

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:23:01 ; Search time 187 Seconds  
(without alignments)  
126.879 Million cell updates/sec

Title: US-10-644-927-1  
Perfect score: 304  
Sequence: 1 KTYGNGVHCTKNSLWKV.....GRLDILLGWATGAFGKTFH 54

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_21.\*  
1: Geneseq1980s.\*  
2: Geneseq1990s.\*  
3: Geneseq2000s.\*  
4: Geneseq2001s.\*  
5: Geneseq2002s.\*  
6: Geneseq2003as.\*  
7: Geneseq2003bs.\*  
8: Geneseq2004s.\*  
9: Geneseq2005s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	304	100.0	54	9 AEB18127	Aeb18127 Lactobaci
2	69.5	22.9	420	6 ADA33974	Ada33974 Acinetoba
3	64	21.1	61	2 AAW17990	Aaw17990 Sakacin P
4	62.5	20.6	558	8 ADR89438	Adr89438 crybun2-o
5	62	20.4	42	2 AAW11633	Aaw11633 Lactobaci
6	60	19.7	37	2 AAW66446	Aaw66446 Cationic
7	60	19.7	37	3 AAY91745	Aay91745 Cationic
8	60	19.7	300	2 AAR56481	Aar56481 CD38. 3/2
9	60	19.7	300	2 AAW37928	Aaw37928 Amino aci
10	60	19.7	300	4 AAB69069	Aab69069 Human CD3
11	60	19.7	300	5 ABG61817	Abg61817 Prostate
12	60	19.7	300	6 ABU09711	Abu09711 CD38 prot
13	60	19.7	300	6 ABU09714	Abu09714 CD38 prot
14	60	19.7	300	7 ADD18904	Add18904 Human dis
15	60	19.7	300	7 ADD45262	Add45262 Human Pro
16	60	19.7	300	7 ADN40030	Adn40030 Cancer/an
17	60	19.7	300	8 ADM32177	Adm32177 Human CD3
18	60	19.7	300	8 ADQ59526	Adq59526 Human can
19	60	19.7	300	8 ADQ88192	Adq88192 Human 903
20	60	19.7	300	9 ADY15608	Ady15608 PRO polyp
21	60	19.7	300	9 ADY19608	Ady19608 PRO polyp
22	60	19.7	300	9 AD213907	Ad213907 Human can
23	60	19.7	688	4 AEG06883	Aeg06883 Novel hum
24	59.5	19.6	513	2 AAW79755	Aaw79755 Euphorbia

25	59	19.4	44	2 AAR14564	Aar14564 Bacterioc
26	59	19.4	44	2 AAR91745	Aar91745 Brevicin,
27	59	19.4	62	2 AAR25078	Aar25078 PA-1 bact
28	59	19.4	300	6 ABU09712	Abu09712 CD38 prot
29	59	19.4	399	8 ADN23488	Adn23488 Bacterial
30	59	19.4	753	4 AAB50093	Aab50093 GB1 prote
31	57.5	18.9	461	7 ADC64559	Adc64559 Trichodes
32	57.5	18.9	461	8 ADQ07360	Adq07360 Trichodes
33	57.5	18.9	461	9 AEA00462	Aea00462 Inorganic
34	57.5	18.9	643	4 ABB71088	Abb71088 Drosophil
35	57	18.8	581	6 ABU44608	Abu44608 Protein e
36	57	18.8	1565	6 ABM69065	Abm69065 Photorhab
37	56.5	18.6	45	5 ABB05473	Abb05473 Enterococ
38	56.5	18.6	242	7 ADF04643	Adf04643 Bacterial
39	56.5	18.6	527	9 AEB28045	Aeb28045 Bacteriop
40	56	18.4	66	7 ADC95053	Adc95053 E. faeciu
41	56	18.4	6641	6 ABU42656	Abu42656 Protein e
42	56	18.4	10182	5 ABP38314	Abp38314 Staphyloc
43	56	18.4	10203	8 ADO84851	Ado84851 S epiderm
44	56	18.4	10203	8 ADO84803	Ado84803 Staphyloc
45	55.5	18.3	244	2 AAY06359	Aay06359 Fusarium

ALIGNMENTS

RESULT 1

AEB18127	AE	B18127	standard; protein; 54 AA.
XX	AC	AE	B18127;
XX	XX		
DT	22-SEP-2005	(first entry)	
XX			
DE			Lactobacillus salivarius strain PVD32 bacteriocin, OR7, SEQ ID NO.1.
XX			
KW			Bacteriocin; feedstuff; therapeutic; gene therapy; lactocin.
XX			
OS			Lactobacillus salivarius; strain PVD32.
XX			
PN			US2005153881-A1.
XX			
PD			14-JUL-2005.
XX			
PF			21-AUG-2003; 2003US-00644927.
XX			
PR			21-AUG-2003; 2003US-00644927.
XX			
PA	(STER/)	STERN N J.	
PA	(SVET/)	SVETOCH E A.	
PA	(ERUS/)	ERUSLANOV B V.	
PA	(VOLO/)	VOLODINA L I.	
PA	(KOVA/)	KOVALEV Y N.	
PA	(KUDR/)	KUDRYAVTSEVA T Y.	
PA	(PERE/)	PERELYGIN V V.	
PA	(POKH/)	POKHILENKO V D.	
PA	(LEVCH/)	LEVCHUK V P.	
PA	(BORZ/)	BORZENKOV V N.	
PA	(SVET/)	SVETOCH O E.	
PA	(MITS/)	MITSEVICH E V.	
PA	(MITS/)	MITSEVICH I P.	
XX			Stern NJ, Svetoch EA, Eruslanov BV, Volodina LI, Kovalev YN;
PI			Kudryavtseva TY, Perelygin VV, Pokhilenko VD, Levchuk VP;
PI			Borzenkov VN, Svetoch OE, Mitsevich EV, Mitsevich IP;
XX			WPI; 2005-496854/50.
XX			
PT			New bacteriocin produced by a lactic acid producing bacterial strain
PT			having the characteristics of strain NRRL B-30514, B-30510, B-30511 or B-
PT			30645, useful in producing therapeutic feed against bacterial
PT			colonization in poultry.
XX			

PS Claim 2; SEQ ID NO 1; 23pp; English.

XX The invention relates to bacteriocins produced by a lactic acid producing

CC bacterial strain having the characteristics of strain NRRL B-30514, B-

CC 30510, B-30511 or B-30645. The bacteriocin is useful in producing a

CC therapeutic feed for animals for reducing the level of colonization by at

CC least one target bacteria in animals, e.g., in poultry and in gene

CC therapy. The present sequence is the Lactobacillus salivarius strain

CC PVD32 (NRRL B-30514) bacteriocin (lactocin) protein.

XX

SQ Sequence 54 AA;

Query Match 100.0%; Score 304; DB 9; Length 54;

Best Local Similarity 100.0%; Pred. No. 2.6e-31;

Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 KTYGTNGVHCTKNSLWGKVRLLKMKYDQNTTYMGRLLQDILLGWATGAFGKTFH 54

DB 1 KTYGTNGVHCTKNSLWGKVRLLKMKYDQNTTYMGRLLQDILLGWATGAFGKTFH 54

RESULT 2

ADA33974

ID ADA33974 standard; protein; 420 AA.

AC ADA33974;

XX

DT 20-NOV-2003 (first entry)

XX

DE Acinetobacter baumannii protein #1135.

XX

KW Acinetobacter baumannii; bacterial disease; antibacterial; vaccine;

KM plant biocontrol agent.

XX

OS Acinetobacter baumannii.

XX

PN US6562958-B1.

XX

PD 13-MAY-2003.

XX

PF 04-JUN-1999; 99US-00328352.

XX

PR 09-JUN-1998; 98US-0088701P.

XX

PA (GENO-) GENOME THERAPEUTICS CORP.

XX

PI Breton G, Bush D;

XX

DR WPI; 2003-576092/54.

DR N-PSDB; ADA29848.

XX

XX New Acinetobacter baumannii proteins and nucleic acids, useful as reagents

PT for diagnosing a bacterial disease, as components of antibacterial

PT vaccines, as targets for antibacterial drugs, or as biocontrol agents for

PT plants.

XX

PS Example; SEQ ID NO 5261; 328pp; English.

XX

CC The invention relates to isolated Acinetobacter baumannii nucleic acids.

CC The A. baumannii nucleic acids and polypeptides are useful as reagents

CC for diagnosing a bacterial disease, as components of antibacterial

CC vaccines, as targets for antibacterial drugs, to detect the presence of

CC A. baumannii and other Acinetobacter species in a sample, in screening

CC compounds for the ability to interfere with the A. baumannii life cycle

CC or to inhibit A. baumannii infection, and as biocontrol agents for

CC plants. The present sequence represents the amino acid sequence of an A.

CC baumannii protein.

XX

SQ Sequence 420 AA;

Query Match 22.9%; Score 69.5; DB 6; Length 420;

Best Local Similarity 38.5%; Pred. No. 4.4;

Matches 20; Conservative 8; Mismatches 13; Indels 11; Gaps 4;

QY 1 KTYGTN---GVHCTKNSLWGKVRLLKMKY---DQNTTYMGRLLQDILLGW 44

DB 361 ESYVTNPXSGKHSNKANIWK---NNTPYSGDRDDTYTTRRLRDTSPILGW 409

RESULT 3

AAW17990

ID AAW17990 standard; protein; 61 AA.

XX

AC AAW17990;

XX

DT 17-OCT-2003 (revised)

DT 10-DEC-1997 (first entry)

XX

DE Sakacin P.

XX

KW IF gene; gene expression; promoter; lactic acid bacterium; bacteriocin;

KM vaccine; vector; sakacin P; ss.

XX

OS Lactobacillus sakei; strain LTH 673.

XX

FH Key Location/Qualifiers

FT Peptide 1..18

FT /label= Sig\_peptide

FT Protein 19..61

FT /label= Sakacin-P

XX

PN WO9718316-A1.

XX

PD 22-MAY-1997.

XX

PF 13-NOV-1996; 96WO-NO000266.

XX

PR 13-NOV-1995; 95NO-00004575.

XX

PA (EJSJ/) EJSINK V G H.

PA (NESI/) NES I F.

PA (BRUR/) BRURBERG M B.

XX

PI Eijsink VGH, Nes IF, Brurberg MB;

XX

DR WPI; 1997-289292/26.

DR N-PSDB; AAT67149.

XX

PT Gene expression system providing regulated lactic acid bacteria protein

PT production - uses new inducing peptide involved in bacteriocin synthesis,

PT useful in fermentation and as a drug delivery system.

XX

PS Example 4; Fig 3; 39pp; English.

XX

CC This polypeptide comprises the sakacin C bacteriocin of Lactobacillus

CC sake LTH673 encoded by gene P (see AAT67149). The invention relates to

CC the discovery of a new regulatory mechanism for gene expression in lactic

CC acid bacteria that includes novel, strongly regulatable promoter elements

CC (see AAT67142-48), such as that of the P gene. Expression of genes under

CC the control of such promoters can be induced by addition of a bacteriocin

CC - inducing peptide such as the claimed peptide of L. sake (see AAW17988)

CC or Lactobacillus plantarum (see AAW17989). Typical applications are in

CC fermentations (e.g. where the gene for an enzyme is regulated) and

CC production of specific proteins, or where the bacteria express a surface

CC antigen, as vaccines. (Updated on 17-OCT-2003 to standardise OS field)

XX

SQ Sequence 61 AA;

Query Match 21.1%; Score 64; DB 2; Length 61;

Best Local Similarity 36.7%; Pred. No. 2.2;

Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;

QY 3 YGTNGVHCTKNSL---WGKVRLLKMKYDQNTTYMGRLLQDILLGWATG 47

DB 20 YYG-NGVHCGKHSCTVDWG-----TAIGNIGNNAAANWATG 54



```

DE XX Cationic peptide leukocin A-val 187.
KW XX Indolicidin analogue; resistance; cationic peptide; antibiotic;
KW XX bacterial infection; tolerance; antibacterial; microorganism; bacteria;
KW XX fungus; parasite; virus.
XX OS Leuconostoc gelidum.
XX XX WO9840401-A2.
XX XX 17-SEP-1998.
XX XX
XX XX 10-MAR-1998; 98WO-CA000190.
XX XX 10-MAR-1997; 97US-0040649P.
XX XX 20-AUG-1997; 97US-00915314.
XX XX 26-SEP-1997; 97US-0060099P.
XX XX 25-FEB-1998; 98US-00030619.
XX XX (MICR-) MICROLOGIX BIOTECH INC.
XX XX
XX XX Fraser JR, West MHP, Mcnicol PU;
XX XX
XX XX WPI; 1998-520800/44.
XX XX
XX XX New indolicidin peptide analogues - useful for, e.g. enhancing activity
XX XX of antibiotic or overcoming tolerance, acquired resistance or inherent
XX XX resistance of microorganisms.
XX XX
XX XX Disclosure; Page 11; 105pp; English.
XX XX
XX CC AAM66393 to AAM66469 represent native cationic peptides from the present
XX CC invention. The present invention describes compositions and methods for
XX CC treating infection, especially bacterial infections. The compositions and
XX CC methods use cationic peptides in combination with an antibiotic agent
XX CC which are then administered to a patient to enhance the activity of the
XX CC antibiotic agent, to overcome: (a) tolerance; (b) acquired resistance;
XX CC and (c) inherent resistance. The combinations of antibiotics and cationic
XX CC peptides can provide synergistic activity against a microorganism that is
XX CC tolerant, inherently resistant, or has acquired resistance to an
XX CC antibiotic agent. They can be used for killing e.g. bacteria, fungi,
XX CC parasites and viruses
XX XX
XX XX Sequence 37 AA;
XX XX
XX XX Query Match 19.7%; Score 60; DB 2; Length 37;
XX XX Best Local Similarity 48.4%; Pred. No. 4;
XX XX Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;
XX XX
XX QY 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
XX XX ||| ||||| |||
XX XX 2 YYG-NGVHCTKSGSVNWGEAFSAGVHRLAN 31
XX XX
XX XX RESULT 7
XX XX AAY91745
XX XX ID AAY91745 standard; peptide; 37 AA.
XX XX
XX XX AC
XX XX AAY91745;
XX XX
XX XX 06-JUN-2000 (first entry)
XX XX
XX XX Cationic peptide Leukocin A-val 187 amino acid sequence.
XX XX
XX XX Cationic peptide; tumour; pharmaceutical composition; cancer; treatment;
XX XX leukaemia; polyoxyalkylene-modified; APO; lymphoma; multiple myeloma;
XX XX breast; lung; ovary; cervix; uterus; skin; prostate; liver; colon;
XX XX multidrug resistance.
XX XX
XX XX Unidentified.
XX XX
XX XX WO9965506-A2.
XX XX
DE XX

PD XX 23-DEC-1999.
XX XX
XX XX 14-JUN-1999; 99WO-CA000552.
XX XX
XX XX 12-JUN-1998; 98US-00096541.
XX XX
XX XX (MICR-) MICROLOGIX BIOTECH INC.
XX XX
XX XX Friedland HD, Krieger TU, Taylor R, Erfle D, Fraser JR, West MHP;
XX XX WPI; 2000-223549/19.
XX XX
XX XX Novel pharmaceutical composition containing optionally activated
XX XX polyoxyalkylene-modified cationic peptides, useful for treating tumors.
XX XX
XX XX Disclosure; Page 11; 94pp; English.
XX XX
XX CC This sequence represents a cationic peptide amino acid sequence, which
XX CC can be used in the pharmaceutical composition of the invention. The
XX CC invention relates to a pharmaceutical composition containing at least one
XX CC activated polyoxyalkylene (APO)-modified cationic peptide. The
XX CC modification of peptides with APO increases their activity against tumour
XX CC cells, including those with a multidrug resistant phenotype. The
XX CC pharmaceutical composition can be used to treat tumours, specifically
XX CC lymphoma, leukaemia, multiple myeloma, or tumours of breast, lung, ovary,
XX CC cervix, uterus, skin, prostate, liver and colon
XX XX
XX XX Sequence 37 AA;
XX XX
XX XX Query Match 19.7%; Score 60; DB 3; Length 37;
XX XX Best Local Similarity 48.4%; Pred. No. 4;
XX XX Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;
XX XX
XX QY 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24
XX XX ||| ||||| |||
XX XX 2 YYG-NGVHCTKSGSVNWGEAFSAGVHRLAN 31
XX XX
XX XX RESULT 8
XX XX AAR56481
XX XX ID AAR56481 standard; protein; 300 AA.
XX XX
XX XX AC AAR56481;
XX XX
XX XX 25-MAR-2003 (revised)
XX XX DT 22-FEB-1995 (first entry)
XX XX
XX XX CD38.
XX XX
XX XX CD38; therapeutic; diagnostic.
XX XX
XX XX Homo sapiens.
XX XX
XX XX WO9417184-A1.
XX XX
XX XX 04-AUG-1994.
XX XX
XX XX 27-JAN-1994; 94WO-US000517.
XX XX
XX XX 29-JAN-1993; 93US-00010905.
XX XX
XX XX (SCHE ) SCHERING CORP.
XX XX (STRD ) UNIV LELAND STANFORD JUNIOR.
XX XX
XX XX Parkhouse RME, Santos-Argumedo L, Grimaldi JC, Bazan JF, Heath A;
XX XX PI Howard MC, Goodnow CC;
XX XX WPI; 1994-264098/32.
XX XX DR N-PSDB; AAQ70644.
XX XX
XX XX Modulation of responses of lymphocytes - using antibody to CD38, a
XX XX soluble CD38 fragment or a modulator of ADP-ribosyl cyclase or cyclic ADP
XX XX -ribosyl hydrolase.

```

XX	PS	Disclosure; Page 46; 54pp; English.
XX	AC	
CC	CC	The protein is a human CD38 molecule which is used to modulate the physiological response of a lymphocyte, which is useful for the stimulation/inhibition of lymphocyte growth or differentiation, particularly for the establishment of antigen tolerance. (Updated on 25-MAR-2003 to correct PN field.)
CC	CC	
CC	CC	
CC	CC	
XX	XX	Sequence 300 AA;
SQ	SQ	
		Query Match 19.7%; Score 60; DB 2; Length 300;
		Best Local Similarity 35.0%; Pred. No. 49;
		Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Qy	5	GTNGVHCTKNSLWGKVRKLNKMKYDQNTTYMGRLODILGW 44
Db	113	GTQVPCNKILLWSRIKDLAHQFTQVQRDMFTLEDTLGY 152
RESULT 9		
AAW37928		
ID	AAW37928	standard; protein; 300 AA.
XX	AC	
AC	AAW37928;	
XX	XX	
DT	09-SEP-1998	(first entry)
XX	XX	
DE	XX	Amino acid sequence of CD38.
XX	XX	
KW	CD38;	CD38 peptide fragment; anti-CD38 autoantibody; autoimmune disease;
KW	insulin dependent diabetes mellitus;	type II diabetes.
XX	OS	Mammalia.
XX	XX	
PN	WO9816245-A1.	
XX	XX	
PD	23-APR-1998.	
XX	XX	
PF	11-APR-1997;	97WO-JP001259.
XX	XX	
PR	15-OCT-1996;	96JP-00272537.
XX	XX	
PA	(SHIO )	SHIONOGI & CO LTD.
PI	Taminato T;	
XX	XX	
DR	WPI; 1998-251054/22.	
DR	N-PSDB; AAV29155.	
XX	XX	
PT	Assaying anti-CD38 auto-antibody - useful for detecting auto-immune disease, e.g. type II diabetes.	
XX	XX	
PS	Disclosure; Page 8-10; 17pp; Japanese.	
XX	XX	
CC	CC	This is the amino acid sequence of the CD38 protein, of which a fragment is used in the method of the invention to create the anti-CD38 autoantibody. The method is used for detecting autoimmune disease, e.g. insulin dependent diabetes mellitus or type II diabetes
XX	XX	
SQ	SQ	Sequence 300 AA;
		Query Match 19.7%; Score 60; DB 2; Length 300;
		Best Local Similarity 35.0%; Pred. No. 49;
		Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Qy	5	GTNGVHCTKNSLWGKVRKLNKMKYDQNTTYMGRLODILGW 44
Db	113	GTQVPCNKILLWSRIKDLAHQFTQVQRDMFTLEDTLGY 152
RESULT 10		
AAB69069		
ID	AAB69069	standard; protein; 300 AA.
XX	AC	
AC	AAB69069;	
XX	XX	
DT	19-APR-2001	(first entry)
XX	XX	
DE	Human CD38	protein sequence SEQ ID NO:2.
XX	XX	
KW	Human; CD38;	diabetes mellitus; detection; cyclic ADP-ribose; cADPR.
XX	XX	
OS	XX	Homo sapiens.
XX	XX	
PN	JP2000316578-A.	
XX	XX	
PD	21-NOV-2000.	
XX	XX	
PF	12-MAY-1999;	99JP-00131955.
XX	XX	
PR	12-MAY-1999;	99JP-00131955.
XX	XX	
PA	(BMLB-) BML KK.	
PA	(KANE/) KANETSUKA A.	
PA	(OKAW/) OKAMOTO H.	
XX	XX	
DR	WPI; 2001-128255/14.	
DR	N-PSDB; AAF32487.	
XX	XX	
PT	Detecting onset of diabetes mellitus comprises detecting specific gene mutations in the CD38 gene.	
XX	XX	
PS	Example; Page 11-12; 19pp; Japanese.	
XX	XX	
CC	CC	The present invention describes a method using a mutation in the CD38 gene (involved in the production of cyclic ADP-ribose (cADPR)), to detect the onset of diabetes mellitus. The method is useful for detecting the onset of diabetes mellitus. The present sequence represents human CD38, which is used in an example from the present invention
XX	XX	
SQ	SQ	Sequence 300 AA;
		Query Match 19.7%; Score 60; DB 4; Length 300;
		Best Local Similarity 35.0%; Pred. No. 49;
		Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Qy	5	GTNGVHCTKNSLWGKVRKLNKMKYDQNTTYMGRLODILGW 44
Db	113	GTQVPCNKILLWSRIKDLAHQFTQVQRDMFTLEDTLGY 152
RESULT 11		
ABG61817		
ID	ABG61817	standard; protein; 300 AA.
XX	AC	
AC	ABG61817;	
XX	XX	
DT	15-AUG-2002	(first entry)
XX	XX	
DE	Prostate cancer-associated protein #18.	
XX	XX	
KW	Prostate cancer;	prostate tumour tissue; human; mammal; cytostatic.
XX	XX	
OS	XX	Mammalia.
XX	XX	
PN	WO200230268-A2.	
XX	XX	
PD	18-APR-2002.	
XX	XX	
PF	12-OCT-2001;	2001WO-US032045.
XX	XX	
PR	13-OCT-2000;	2000US-00687576.
PR	08-DEC-2000;	2000US-00733288.
PR	08-DEC-2000;	2000US-00733742.
PR	24-JAN-2001;	2001US-0263957P.

```

PR 16-MAR-2001; 2001US-0276791P.
PR 16-MAR-2001; 2001US-0276888P.
PR 06-APR-2001; 2001US-0281922P.
PR 24-APR-2001; 2001US-0286214P.
PR 30-APR-2001; 2001US-00847046.
PR 04-MAY-2001; 2001US-0288589P.
XX (EOSB-) EOS BIOTECHNOLOGY INC.
PA
XX Gish KC, Mack DH, Wilson KE, Afar D, Hevezi P;
XX WPI; 2002-471335/50.
XX N-PSDB; ABK92132.
XX
PT Detecting a prostate cancer-associated transcript in a cell in a patient,
PT useful for diagnosing prostate cancer (PC) or screening modulators of PC,
PT by determining if prostate cancer-associated genes are expressed in a
PT prostate tissue.
XX
XX Claim 27; Page 314; 436pp; English.
XX
CC The present invention relates to methods of detecting a prostate cancer-
CC associated transcript in a cell from a patient. The method comprises
CC contacting a biological sample from the patient with prostate cancer-
CC associated polynucleotides (designated PC genes) that selectively
CC hybridise to a sequence that is at least 80% identical to them. The
CC prostate cancer-associated polynucleotide sequences are differentially
CC expressed in prostate tumour tissue or in prostate cancer and are derived
CC from the tissues of various organisms such as humans or other mammals
CC (e.g. mice, sheep and dogs). The methods of the invention are useful for
CC diagnosing and treating prostate cancer in mammals. The prostate cancer-
CC associated genes are useful for diagnosing or treating prostate cancer,
CC as well as for identifying modulators of prostate cancer or agents that
CC inhibit prostate cancer. The nucleic acid sequences are particularly
CC useful in gene therapy, as a vaccine or in antisense applications.
CC ABG61800-ABG61944 represent prostate cancer-associated proteins
XX
XX SQ Sequence 300 AA;
XX
Query Match 19.7%; Score 60; DB 5; Length 300;
Best Local Similarity 35.0%; Pred. No. 49;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Oy 5 GTNGVCHTNSLWGKVKRLKNMKYDQNTYMGRLQDILLGW 44
Db 113 GTQVPCNKILLWSRIKDLAQHTQVQVRDMFTLEDTLGY 152
||| ||| ||| ::| ||| ||| |||
113 GTQVPCNKILLWSRIKDLAQHTQVQVRDMFTLEDTLGY 152

RESULT 12
ABU09711
ID ABU09711 standard; protein; 300 AA.
XX
XX AC ABU09711;
XX
XX DT 03-JUL-2003 (first entry)
XX
XX DE CD38 protein.
XX
XX KW CD38; diabetic onset; diabetes; lymphocyte surface marker.
XX
XX OS Hominidae.
XX
XX FH Key Location/Qualifiers
XX Misc-difference 264
XX FT /note= "Wild type Ser substituted by Leu"
XX
XX PN US2003027134-A1.
XX
XX PD 06-FEB-2003.
XX
XX PF 31-JAN-2001; 2001US-00773307.
XX
XX PR 31-JAN-2001; 2001US-00773307.
XX
XX PA (EGAS/) EGASHIRA T.
XX PA (NAGA/) NAGANO M.
XX PA (SAGE/) SAGEHASHI Y.
XX PA (MATS/) MATSUI K.
XX PA (HATT/) HATTORI H.
XX PA (KANA/) KANATSUKA A.
XX PA (TAKA/) TAKASAWA S.
XX PA (OKAM/) OKAMOTO H.
XX
XX PI Egashira T, Nagano M, Sagehashi Y, Matsui K, Hattori H;

PA (MATS/) MATSUI K.
PA (HATT/) HATTORI H.
PA (KANA/) KANATSUKA A.
PA (TAKA/) TAKASAWA S.
PA (OKAM/) OKAMOTO H.
XX
XX PI Egashira T, Nagano M, Sagehashi Y, Matsui K, Hattori H;
XX
WPI; 2003-417248/39.
N-PSDB; ACA60951.
XX
DR DR
XX
XX PT Detecting risk factor for onset of diabetes in an individual, involves
XX detecting genetic abnormality of the gene CD38.
XX
XX PS Disclosure; Page 10; 29pp; English.
XX
XX CC The invention describes a method of detecting a risk factor for diabetic
XX onset in an individual. The method involves detecting genetic abnormality
XX of the gene CD38 which is a human lymphocyte surface marker and is a risk
XX gene for onset of diabetes. The sites of abnormality in CD38 gene include
XX CC a site encoding Arg at residue 140, a site encoding Ser at residue 264,
XX CC or guanine at nucleotide position -28 in intron 7. This is the amino acid
XX sequence of wild type CD38 protein
XX
XX SQ Sequence 300 AA;
XX
Query Match 19.7%; Score 60; DB 6; Length 300;
Best Local Similarity 35.0%; Pred. No. 49;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;
Oy 5 GTNGVCHTNSLWGKVKRLKNMKYDQNTYMGRLQDILLGW 44
Db 113 GTQVPCNKILLWSRIKDLAQHTQVQVRDMFTLEDTLGY 152
||| ||| ||| ::| ||| ||| |||
113 GTQVPCNKILLWSRIKDLAQHTQVQVRDMFTLEDTLGY 152

RESULT 13
ABU09714
ID ABU09714 standard; protein; 300 AA.
XX
XX AC ABU09714;
XX
XX DT 03-JUL-2003 (first entry)
XX
XX DE CD38 protein S264L mutant.
XX
XX KW CD38; diabetic onset; diabetes; lymphocyte surface marker; mutant;
XX mutein.
XX
XX OS Hominidae.
XX
XX FH Key Location/Qualifiers
XX Misc-difference 264
XX FT /note= "Wild type Ser substituted by Leu"
XX
XX PN US2003027134-A1.
XX
XX PD 06-FEB-2003.
XX
XX PF 31-JAN-2001; 2001US-00773307.
XX
XX PR 31-JAN-2001; 2001US-00773307.
XX
XX PA (EGAS/) EGASHIRA T.
XX PA (NAGA/) NAGANO M.
XX PA (SAGE/) SAGEHASHI Y.
XX PA (MATS/) MATSUI K.
XX PA (HATT/) HATTORI H.
XX PA (KANA/) KANATSUKA A.
XX PA (TAKA/) TAKASAWA S.
XX PA (OKAM/) OKAMOTO H.
XX
XX PI Egashira T, Nagano M, Sagehashi Y, Matsui K, Hattori H;

```

XX  
Sd

CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 300 AA;

Query Match 19.7%; Score 60; DB 7; Length 300;  
Best Local Similarity 35.0%; Pred. No. 49;  
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;  
Qy 5 GTNGVHCTKNSLWGKVRKKNMKYDQNTTYMGRLODILLGW 44  
Db 113 GTQVPCNKILLWSRIKDLAHQFTQVQDRDMFTLEDTLGY 152

Search completed: May 2, 2006, 17:26:33  
Job time : 191 secs



GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:26:51 ; Search time 38 Seconds  
(without alignments)  
136.729 Million cell updates/sec

Title: US-10-644-927-1  
Perfect score: 304  
Sequence: 1 KTYVGTNGVHCTKNSLWGVK.....GRLODILLGWATGAFGKTFH 54

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_80.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65.5	21.5	471	2 S11899	glutamate-ammonia
2	65.5	21.5	471	2 A99180	hypothetical prote
3	64	21.1	61	1 S38508	sakacin P precursor
4	63	20.7	590	2 AF5555	hypothetical prote
5	62.5	20.6	446	2 A43995	glutamate-ammonia
6	60	19.7	61	2 A41657	leucocin A-UAL 187
7	60	19.7	61	2 S52208	mesentericin Y105
8	60	19.7	300	2 A43521	lymphocyte surface
9	59.5	19.6	286	2 F84954	spermidine synthas
10	59	19.4	62	1 A48941	pediocin PA-1 prec
11	59	19.4	399	2 T20455	hypothetical prote
12	58	19.1	3036	2 T18995	hypothetical prote
13	57.5	18.9	510	2 T50021	inositol-3-phospha
14	56.5	18.6	517	1 G18PT4	gene 12 protein -
15	55.5	18.3	132	2 PC3131	hepatocyte growth
16	55.5	18.3	934	2 B28838	parasporal crystal
17	55.5	18.3	1332	2 F69732	PBSX prophage ORF
18	55	18.1	201	2 AD3302	31K outer-membrane
19	55	18.1	346	2 B90243	aminomethyltransfe
20	55	18.1	369	2 T24205	hypothetical prote
21	55	18.1	516	2 G84442	probable nucleosid
22	55	18.1	799	1 TVRTTB	nerve growth facto
23	54.5	17.9	319	2 T15463	hypothetical prote
24	54.5	17.9	322	2 D86760	phosphate starvati
25	54.5	17.9	374	2 F97257	glycosyltransferas
26	54.5	17.9	383	2 B96806	hypothetical prote
27	54.5	17.9	494	1 B41141	1-aminocyclopropan
28	54	17.8	262	1 TLBPM1	tail fiber protein
29	54	17.8	269	2 S51815	hypothetical prote

30	54	17.8	494	2 JC2382	sodium/proline sym
31	53.5	17.6	254	2 S31393	chlorophyll a/b-b1
32	53.5	17.6	324	2 G69515	transcription regu
33	53.5	17.6	462	1 YSBVTM	threonine-tRNA lig
34	53.5	17.6	510	2 D84610	probable myo-inosi
35	53.5	17.6	527	2 S29530	gene 12 protein
36	53.5	17.6	1091	1 PL0009	complement C3d/Eps
37	53.5	17.6	1166	2 H71609	hypothetical prote
38	53.5	17.6	1585	2 B69948	phage-related prot
39	53	17.4	167	2 S49628	FUN81 protein - ye
40	53	17.4	202	2 C90228	hypothetical prote
41	53	17.4	249	2 B95357	probable EtfB2 ele
42	53	17.4	258	2 T13591	tail fiber adhesin
43	53	17.4	264	2 JU0328	cellulase (EC 3.2.
44	53	17.4	327	2 AI3308	glutathione transf
45	53	17.4	327	2 S54560	TOM37 protein - ye

ALIGNMENTS

RESULT 1

S11899  
glutamate-ammonia ligase (EC 6.3.1.2) - Sulfolobus solfataricus  
N:Alternate names: Glutamine synthetase  
C:Species: Sulfolobus solfataricus  
C>Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C:Accession: S11899  
R:Sanangelantoni, A.M.; Barbarini, D.; di Pasquale, G.; Cammarano, P.; Tiboni, O.  
Mol. Gen. Genet. 221, 187-194, 1990  
A:Title: Cloning and nucleotide sequence of an archaeobacterial glutamine synthetase gene  
A:Reference number: S11899; MUID:90318316; PMID:1973523  
A:Accession: S11899  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-471 <SAN>  
A:Cross-references: UNIPROT:P23794; UNIPARC:UPI0000170275; GB:X53263; NID:G2968664; PIDN  
C:Superfamily: glutamate-ammonia ligase  
C:Keywords: ligase

Query Match 21.5%; Score 65.5; DB 2; Length 471;  
Best Local Similarity 48.6%; Pred. No. 2;  
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;

QY 1 KTYVGTN--GVHCTKNSLWGVKVRUKMKYQDNTTY 33  
DB 255 KPIYGDNGTGMH-TLHSLWTKDGKKNLMYDPNDEY 288

RESULT 2

A99180  
hypothetical protein glna-1 [imported] - Sulfolobus solfataricus  
C:Species: Sulfolobus solfataricus  
C>Date: 24-May-2001 #sequence\_revision 24-May-2001 #text\_change 09-Jul-2004  
C:Accession: A99180  
R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan  
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, J.  
arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.  
submitted to GenBank, April 2001  
A:Description: Sulfolobus solfataricus complete genome.  
A:Reference number: A99139  
A:Accession: A99180  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-471 <KUR>  
A:Cross-references: UNIPROT:P23794; UNIPARC:UPI000012B704; GB:AE0066641; NID:gl3813510;  
C:Genetics:  
A:Gene: glna-1  
C:Superfamily: glutamate-ammonia ligase

Query Match 21.5%; Score 65.5; DB 2; Length 471;  
Best Local Similarity 48.6%; Pred. No. 2;  
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;





F;19-62/Product: peditocin PA-1 #status experimental <MAT>  
F;27-32,42-62/Disulfide bonds: #status experimental

Query Match 19.4%; Score 59; DB 1; Length 62;  
Best Local Similarity 37.5%; Pred. No. 1.4;  
Matches 18; Conservative 5; Mismatches 9; Indels 16; Gaps 4;

QY 3 YYGTNGVHCTKNSL---WGKVLKMKYDQNTTYMGRLODILLGWATG 47  
||| ||| ||| ||| ||| ||| : : : |||  
Db 20 YYG-NGVTGKHGSCVDWGKA-----TTCI--INNGAMAWATG 54

## RESULT 11

T20455

hypotheical protein F01D4.3 - Caenorhabditis elegans

A;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 12-Jul-2004

C;Accession: T20455

R;Wild, A.

submitted to the EMBL Data Library, October 1996

A;Reference number: Z19278

A;Accession: T20455

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-399 &lt;WIL&gt;

A;Cross-references: UNIPROT:O17755; UNIPARC:UPI000007E797; EMBL:Z81054; PIDN: CAB02882.1;

A;Experimental source: clone F01D4

C;Genetics:

A;Gene: CESP:F01D4.3

A;Map position: 4

A;Introns: 45/3; 235/2; 294/2

Query Match 19.4%; Score 59; DB 2; Length 399;  
Best Local Similarity 35.8%; Pred. No. 11;  
Matches 19; Conservative 4; Mismatches 22; Indels 8; Gaps 3;

QY 3 YYGTNGVHCTKNSLWGKVLKMKYDQNTTYMGRLODILLG--WATGARGKTF 53  
||| ||| ||| ||| ||| ||| : : : |||  
Db 75 YYTDSGSCSN---GRFQKPNLPLOPWEYMH--SDVTAGKVLGEGAFGRVF 121

## RESULT 12

T18995

hypotheical protein C0688.7 - Caenorhabditis elegans

A;Species: Caenorhabditis elegans

C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004

C;Accession: T18995

R;Steward, C.

submitted to the EMBL Data Library, November 1996

A;Reference number: Z19057

A;Accession: T18995

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-3036 &lt;WIL&gt;

A;Cross-references: UNIPROT:O17575; UNIPARC:UPI000017B740; EMBL:Z81463; PIDN: CAB03852.1;

A;Experimental source: clone C0688

C;Genetics:

A;Gene: CESP:C0688.7

A;Map position: 5

A;Introns: 102/3; 157/2; 288/2; 439/3; 542/1; 591/3; 798/3; 835/3; 1139/1; 1204/3; 1325/

Query Match 19.1%; Score 58; DB 2; Length 3036;  
Best Local Similarity 26.1%; Pred. No. 1.5e+02;  
Matches 18; Conservative 5; Mismatches 24; Indels 22; Gaps 2;

QY 1 KTYVGTGVHCT-----KNSLWGKVLKMKYDQNTTYMGRLL-----Q 38  
||| ||| ||| ||| ||| ||| : : : |||  
Db 543 KAYGLNVTICIEKTCNGVGFANDIRERTALTNTLDENQGVAGFLVKDGAIDLWLNET 602

QY 39 DILLGWATG 47

||| ||| |||

Db 603 RILNNWGDG 611

## RESULT 13

T50021

inositol-3-phosphate synthase (EC 5.5.1.4) T31P16.160 [similarity] - Arabidopsis thaliana

N;Alternate names: protein T31P16.160

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 09-Jul-2004

C;Accession: T50021

R;Bevan, M.; Zimmermann, W.; Grueneisen, A.; Wambutt, R.; Kalicki, J.; Wohlmann, P.; Sm

submitted to the Protein Sequence Database, May 2000

A;Reference number: Z25027

A;Accession: T50021

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-510 &lt;BEV&gt;

A;Cross-references: UNIPROT:Q9LX12; UNIPARC:UPI0000001794; EMBL:AL356332; GSPDB:GN000063;

A;Experimental source: cultivar Columbia; BAC clone T31P16

C;Genetics:

A;Gene: ATSP:T31P16.160

A;Map position: 5

A;Introns: 63/2; 86/2; 131/3; 214/2; 290/1; 328/3; 387/3; 450/3; 471/3

C;Superfamily: myo-inositol-1-phosphate synthase

C;Keywords: intramolecular lyase; isomerase; NAD

Query Match 18.9%; Score 57.5; DB 2; Length 510;  
Best Local Similarity 37.7%; Pred. No. 23;  
Matches 20; Conservative 8; Mismatches 20; Indels 5; Gaps 4;

QY 3 YYGTNGVHCTKNSL--WGKVLKMKYDQNT-TYMGRLQDILLGWATGAFGKT 52  
||| ||| ||| ||| ||| ||| : : : |||  
Db 26 YQTTTELVENKNGAFQW-TVKPKTVKYEKFTDTHVPKLGVMVLGVWG-GNNGST 76

## RESULT 14

GIBPT4

Gene 12 protein - phase T4

N;Alternate names: tail fiber protein

C;Species: phase T4

C;Date: 30-Sep-1991 #sequence\_revision 30-Sep-1991 #text\_change 30-Jun-1993

C;Accession: S01889

R;Selivanov, N.A.; Prilipov, A.G.; Mesyanzhinov, V.V.

Nucleic Acids Res. 16, 2334, 1988

A;Title: Nucleotide and deduced amino acid sequence of bacteriophage T4 gene 12.

A;Reference number: S01889; MUID:88189824; PMID:3357780

A;Accession: S01889

A;Molecule type: DNA

A;Residues: 1-517 &lt;SEL&gt;

A;Cross-references: UNIPARC:UPI0000174BAD; EMBL:X06792

A;Note: the authors translated the codon CAG for residue 279 as His

C;Genetics:

A;Gene: 12

C;Superfamily: phase T4 gene 12 protein

C;Keywords: tail fiber

Query Match 18.6%; Score 56.5; DB 1; Length 517;  
Best Local Similarity 33.3%; Pred. No. 31;  
Matches 15; Conservative 8; Mismatches 15; Indels 7; Gaps 2;

QY 8 GVHCTKNSLWGKVLKMKYDQNTTYMGRLODILLGWATGAFGKT 52  
||| ||| ||| ||| ||| ||| : : : |||  
Db 416 GVGCTGGYV-GEVQIQMSYHKHAGGFGEHDDL-----GAFGNT 453

## RESULT 15

PC2131

hepatocyte growth factor receptor - rat (fragment)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 03-May-1994 #sequence\_revision 07-Oct-1994 #text\_change 05-Oct-2004

C;Accession: PC2131

R;Tsujii, M.; Kawano, S.; Tsuji, S.; Ito, T.; Hayashi, N.; Horimoto, M.; Mita, E.; Nagano

Biochem. Biophys. Res. Commun. 200, 536-541, 1994

A;Title: Increased expression of c-met messenger RNA following acute gastric injury in rat

A;Reference number: PC2131; MUID:94220137; PMID:8166728

Search completed: May 2, 2006, 17:31:16  
Job time : 41 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:23:40 ; Search time 234 Seconds  
(without alignment)  
162.814 Million cell updates/sec

Title: US-10-644-927-1

Perfect score: 304

Sequence: 1 KTYGTNGVHCTKNSLWGVK.....GRLQDILLGWATGAFGKTFH 54

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_05.80.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	195	64.1	81	Q48496_LACAC	Q48496 lactobacill
2	65.5	21.5	471	1 GLNA SULSO	P23794 sulfolobus
3	65	21.4	301	1 CD38_MACFA	Q5van0 macaca falc
4	65	21.4	325	2 Q5XQ46_9PEZI	Q5xq46 cryptovalsa
5	64.5	21.2	460	2 Q4NH89_9MITCC	Q4nh89 arthrobacte
6	64.5	21.2	1071	2 Q75CA6_ASHGO	Q75ca6 ashbya gos
7	64	21.1	61	1 SAKP_LACSK	P35618 lactobacill
8	64	21.1	61	2 Q7B5F5_LACSK	Q7b5f5 lactobacill
9	64	21.1	714	2 Q51P75_MAGGR	Q51p75 magnaporth
10	63	20.7	590	2 Q8YK80_ANAEP	Q8yk80 anabaena sp
11	63	20.7	1582	2 Q8RIM1_FUSNN	Q8rim1 fusobacteri
12	63	20.7	1630	2 Q8RHH7_FUSNN	Q8rhh7 fusobacteri
13	62.5	20.6	446	1 GLNA_METVO	P21154 methanococ
14	62.5	20.6	558	2 Q8VM63_BACTA	Q8vm63 bacillus th
15	61.5	20.2	469	2 Q9PZ03_GVXN	Q9pz03 xestia c-ni
16	61.5	20.2	524	2 Q7OED6_ANOGA	Q7qed6 anopheles g
17	61.5	20.2	547	2 Q8A1J6_BACTN	Q8ajj6 bacteroides
18	61.5	20.2	853	2 Q9DL93_9HIV1	Q9dl93 human immun
19	61.5	20.2	858	2 Q9DL79_9HIV1	Q9dl79 human immun
20	61.5	20.2	1474	2 Q4Q2Y5_LEIMA	Q4q2y5 leishmania
21	61	20.1	61	2 Q4U1B4_LACCU	Q4uib4 lactobacill
22	61	20.1	684	2 Q5WI64_LACCU	Q5wi64 bacillus cl
23	61	20.1	1524	2 Q4UIZ1_THEAN	Q4uiz1 theileria a
24	60	19.7	61	1 LCCA_LEUGE	P34034 leuconostoc
25	60	19.7	61	1 LCCA_LEUCA	Q53446 leuconostoc
26	60	19.7	61	1 MTCV_LEUME	P38577 leuconostoc
27	60	19.7	61	2 Q791V9_LEUME	Q791v9 leuconostoc
28	60	19.7	300	1 CD38_HTMAN	P28907 homo sapien
29	60	19.7	498	2 Q622U2_CAEBR	Q622u2 caenorhabdi
30	60	19.7	1143	2 Q948Y6_VOLCA	Q948y6 volvox cart
31	60	19.7	1794	2 Q8RHH1_FUSNN	Q8rhh1 fusobacteri

ALIGNMENTS

RESULT 1

Q48496\_LACAC PRELIMINARY; PRT; 81 AA.  
ID Q48496\_LACAC PRELIMINARY; PRT; 81 AA.  
AC Q48496;  
DT 01-NOV-1996 (Tremblrel. 01, Created)  
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Acidocin A precursor.  
GN Name=acdA;  
OS Lactobacillus acidophilus.  
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;  
OC Lactobacillus.  
OX NCBI\_TaxID=1579;  
RN [1]\_TaxID=1579;  
RP NUCLEOTIDE SEQUENCE.  
RC STRAIN=TK9201;  
RX MEDLINE=95314239; PubMed=7793908;  
RA Kanatani K., Oshimura M., Sano K.;  
RT "Isolation and characterization of acidocin A and cloning of the  
RT bacteriocin gene from Lactobacillus acidophilus.";  
RL Appl. Environ. Microbiol. 61:1061-1067(1995).  
DR EMBL; D37881; BAA07120.1; -; Genomic\_DNA.  
DR InterPro; IPR010133; Bacteriocin\_sig.  
DR TIGRFAMs; TIGR01847; bacteriocin\_sig; 1.  
KW Signal.  
FT SIGNAL 1 23 Potential.  
FT CHAIN 24 81 mature acidocin A.  
SQ SEQUENCE 81 AA; 8946 MW; 89698AA296F7819D CRC64;  
Query Match 64.1%; Score 195; DB 2; Length 81;  
Best Local Similarity 64.6%; Pred. No. 2,3e-17;  
Matches 42; Conservative 1; Mismatches 4; Indels 18; Gaps 3;

QY 1 KTYGTNGVHCTKNSLWGVKRLKNMKYDQNTTYMGRLL-----QD---ILLGWATGAF 49  
Db 24 KTYGTNGVHCTKNSLWGVKRLKNV-----IPGTLCKRQSLPIKQDLKILLGWATGAF 76  
QY 50 GKTFH 54  
Db 77 GKTFH 81

RESULT 2

GLNA\_SULSO STANDARD; PRT; 471 AA.  
ID GLNA\_SULSO STANDARD; PRT; 471 AA.  
AC P23794;  
DT 01-NOV-1991 (Rel. 20, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 10-MAY-2005 (Rel. 47, Last annotation update)  
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase) (GS).  
GN Name=glina; Synonyms=glina-1; OrderedLocusNames=SSO0366;  
OS Sulfolobus solfataricus.  
OC Archaea; Crenarchaeota; Thermoprotei; Sulfolobales; Sulfolobaceae;  
OC Sulfolobus.

```

OX NCBI_TaxID=2287;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RX MEDLINE=90318316; PubMed=1973523; DOI=10.1007/BF00261719;
RA Sanangelantoni A.M., Barbarini D., di Pasquale G., Cammarano P.,
RT Tiboni O.;
RT "Cloning and nucleotide sequence of an archaeobacterial glutamine
RT synthetase gene: phylogenetic implications.";
RL Mol. Gen. Genet. 221:187-194(1990).
[2]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 35092 / DSM 1617 / P2; DOI=10.1073/pnas.141222098;
RX MEDLINE=21332296; PubMed=11427726; DOI=10.1073/pnas.141222098;
RA She Q., Singh R.K., Confalonieri F., Zivanovic Y., Allard G.,
RA Aweys M.J., Chan-Weher C.C.-Y., Clausen I.G., Curtis B.A.,
RA De Moors A., Eraso G., Fletcher C., Gordon P.M.K.,
RA Heikamp-de Jong I., Jeffries A.C., Kozera C.J., Medina N., Peng X.,
RA Thi-Ngoc H.P., Redder P., Schenk M.E., Theriault C., Tolstrup N.,
RA Charlebois R.L., Doolittle W.F., Duguet M., Gaasterland T.,
RA Garrett R.A., Ragan M.A., Sensen C.W., Van der Oost J.;
RT "The complete genome of the crenarchaeon Sulfolobus solfataricus P2.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:7835-7840(2001).
CC -|- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate +
CC L-glutamine.
CC -|- ENZYME REGULATION: The activity of this enzyme is controlled by
CC adenylation under conditions of abundant glutamine. The fully
CC adenylated enzyme complex is inactive (By similarity).
CC -|- SUBCELLULAR LOCATION: Cytoplasmic.
CC -|- SIMILARITY: Belongs to the glutamine synthetase family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; X53263; CAA37353.1; -; Genomic DNA.
DR EMBL; X53263; CAA37352.1; -; Genomic DNA.
DR EMBL; AE006669; AAK40696.1; -; Genomic DNA.
DR PIR; A99180; A99180.
DR PIR; S11899; S11899.
DR HSP; P08201; ILGR.
DR InterPro; IPR008147; Gln_synth_beta.
DR InterPro; IPR008146; Gln_synth_C.
DR InterPro; IPR004809; GlnA.
DR Pfam; PF00120; Gln-synt C; 1.
DR Pfam; PF03951; Gln-synt_N; 1.
DR ProDom; PD001057; Gln_synth_C; 1.
DR TIGRFAMs; TIGR00653; GlnA; 1.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA_ATP; 1.
DR Complete proteome; Ligase.
KW BINDING 400 AMP (covalent) (By similarity).
FT CONFLICT 162 162 S -> T (in Ref. 1).
FT SEQUENCE 471 AA; 53348 MW; 1CC4239287B6C2CD CRC64;
SQ
Query Match 21.5%; Score 65.5; DB 1; Length 471;
Best Local Similarity 48.6%; Pred. No. 11;
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;
OY 1 KTYYGTN--GVHCTKNSLWGVRLKNMKYDQNTTY 33
DB 255 KPIYGDNGTGMH--THLSLWTKDGGKKNLWYDPNDY 288
RESULT 3
CD38 MACFA STANDARD; PRT; 301 AA.
AC Q5VAN0;
DT 10-MAY-2005 (Rel. 47, Created)
DT 10-MAY-2005 (Rel. 47, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE ADP-ribosyl cyclase 1 (EC 3.2.2.5) (Cyclic ADP-ribose hydrolase 1)

```

```

DE (ADPr hydrolase 1) (CD38 homolog).
GN Name=CD38;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
OC Cercopithecoidea; Cercopithecinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP NUCLEOTIDE SEQUENCE [MRNA].
RX PubMed=15383153; DOI=10.1186/1471-2172-5-21;
RA Ferrero E., Orciani M., Vacca P., Ortolan E., Crovella S., Titti F.,
RA Saccucci F., Malavasi F.;
RT "Characterization and phylogenetic epitope mapping of CD38 ADPR
RT cyclase in the cynomolgus macaque.";
RL BMC Immunol. 5:21-21(2004).
CC -|- FUNCTION: Synthesizes cyclic ADP-ribose, a second messenger for
CC glucose-induced insulin secretion. Also has ADPR hydrolase
CC activity. Also moonlights as a receptor in cells of the immune
CC system (By similarity).
CC -|- CATALYTIC ACTIVITY: NAD(+) + H(2)O = ADP-ribose + nicotinamide.
CC -|- ENZYME REGULATION: ATP inhibits the hydrolyzing activity (By
CC similarity).
CC -|- SUBCELLULAR LOCATION: Type II membrane protein (By similarity).
CC -|- SIMILARITY: Belongs to the ADP-ribosyl cyclase family.
-----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
-----
DR EMBL; AY55148; AAT36330.1; -; mRNA.
DR InterPro; IPR003193; Rib_hydrolase.
DR PANTHER; PTHR10912; Rib_hydrolase; 1.
DR Pfam; PF02267; Rib_hydrolase; 1.
KW Glycoprotein; Hydrolase; NAD; Receptor; Signal-anchor; Transmembrane.
FT TOPO_DOM 1 21 Cytoplasmic (Potential).
FT TRANSMEM 22 43 Signal-anchor for type II membrane
FT TOPO_DOM 44 301 Extracellular (Potential).
FT ACT_SITE 120 120 By similarity.
FT ACT_SITE 202 202 By similarity.
FT CARBOHYD 101 101 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 121 121 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 210 210 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 220 220 N-linked (GlcNAc...) (Potential).
FT DISULFID 68 83 By similarity.
FT DISULFID 100 181 By similarity.
FT DISULFID 161 174 By similarity.
FT DISULFID 255 276 By similarity.
FT DISULFID 288 297 By similarity.
SQ SEQUENCE 301 AA; 34422 MW; B659212B926165B1 CRC64;
Query Match 21.4%; Score 65; DB 1; Length 301;
Best Local Similarity 35.0%; Pred. No. 8.1;
Matches 14; Conservative 8; Mismatches 18; Indels 0; Gaps 0;
OY 5 GTNGVHCTKNSLWGVRLKNMKYDQNTTYMGRQLDILLGW 44
DB 114 GTQTPCNKTLMSRIKDLAHQFTQVORDMFTLEDMLGY 153
RESULT 4
Q5XQ46_9PEZI
ID Q5XQ46_9PEZI PRELIMINARY; PRT; 325 AA.
AC Q5XQ46_9PEZI
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Family 10 xylanase (EC 3.2.1.8).
OS Cryptovalsa sp. BCC 7197.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Xylariomycetidae; Xylariales; Diatrypales; Cryptovalsa.

```



```
OX NCBI_TaxID=295079;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=BCC 7197;
RA Boonyapakorn K., Pootanakit K., Chantasingh D., Kirtikara K.,
RA Eurwilaichitr L.;
RT "Cloning and expression of xylanase 10 from Cryptovalsa sp. (BCC7197)
RT in Pichia pastoris.";
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY741212; AAU89274.1; -; mRNA.
DR GO; GO:0031176; F:endo-1,4-beta-xylanase activity; IEA.
DR GO; GO:0005975; P:carbohydrate metabolism; IEA.
DR InterPro; IPR001000; Glyco_hydro_10.
DR Pfam; PF00331; Glyco_hydro_10; 1.
DR PRINTS; PR00134; GLYHDLASE10.
DR SMART; SM00633; Glyco_10; 1.
DR KW Glycosidase; Hydrolase; Xylan degradation.
SQ SEQUENCE 325 AA; 35107 MW; 6E0930A7F8C6433C CRC64;

Query Match 21.4%; Score 65; DB 2; Length 325;
Best Local Similarity 32.8%; Pred. No. 8.8;
Matches 21; Conservative 6; Mismatches 15; Indels 22; Gaps 4;

QY 1 KTYGTGNGVHCTKNSL-----WGKRLKN-MKYDQNTYMGRLQ-----DILL 42
Db 36 KLYTGT-----CTDGLLTSGSAAIIQTNFGQVTPENSMKKWDQTSRGQFNLAQADYLV 91
QY 43 GWAT 46
Db 92 DWAT 95

RESULT 5
Q4NH89 9MICC PRELIMINARY; PRT; 460 AA.
ID Q4NH89 9MICC PRELIMINARY; PRT; 460 AA.
AC Q4NH89
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Putative membrane protein precursor.
GN ORFNames=ArthDRAFT_2550;
OS Arthrobacter sp. FB24
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Micrococccineae; Micrococcaceae; Arthrobacter.
OX NCBI_TaxID=290399;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (JGI-PGF);
RA Copeland A., Lucas S., Lapidus A., Barry K., Detter C., Glavina T.,
RA Hammon N., Izrasi S., Pitluck S., Richardson P.;
RA "Sequencing of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=FB24;
RG US DOE Joint Genome Institute (PGF-ORNL);
RA Larimer F., Land M.;
RA "Annotation of the draft genome assembly of Arthrobacter sp. FB24.";
RL Submitted (JUN-2005) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAHQ01000004; EAL97009.1; -; Genomic_DNA.
KW Signal.
FT SIGNAL 460 460 Potential.
SQ SEQUENCE 460 AA; 49805 MW; F5F6146AFB237113 CRC64;

Query Match 21.2%; Score 64.5; DB 2; Length 460;
Best Local Similarity 27.5%; Pred. No. 15;
Matches 19; Conservative 12; Mismatches 21; Indels 17; Gaps 3;

QY 1 KTYGTGNGVHCTKNSLWGKRL-----KMKYDQNTYMGRLQDILLG-----W 44
```

```
Db 158 RSTFGTSGPHATRNK-WGAVILLAVLNLGFWRETLAFGQINILMGIMADLLARNQRW 216
QY 45 ATGAFGKTF 53
Db 217 NQGFPGRGF 225

RESULT 6
Q75CAG ASHGO PRELIMINARY; PRT; 1071 AA.
ID Q75CAG ASHGO PRELIMINARY; PRT; 1071 AA.
AC Q75CAG;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE ACR006CP.
GN Name=ACR006C;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC STRAIN=ATCC 10895;
RX PubMed=15001715; DOI=10.1126/science.1095781;
RA Dietrich P.S., Voegelé S., Brachat S., Lerch A., Gates K., Steiner S.,
RA Mohr C., Poehlmann R., Luedi P., Choi S., Wing R.A., Flavier A.,
RA Gaffney T.D., Philippsen P.;
RT "The Ashbya gossypii genome as a tool for mapping the ancient
RT Saccharomyces cerevisiae genome.";
RL Science 304:304-307(2004).
DR EMBL; AE016816; AAS51233.1; -; Genomic_DNA.
DR AGP; ACR006C; -;
DR Complete proteome.
SQ SEQUENCE 1071 AA; 123573 MW; 0816B52D60032663 CRC64;

Query Match 21.2%; Score 64.5; DB 2; Length 1071;
Best Local Similarity 31.7%; Pred. No. 39;
Matches 13; Conservative 8; Mismatches 11; Indels 9; Gaps 1;

QY 8 GVHCTKNSLWGKRLKNMKYDQNTYMGRLQDILLGWATGA 48
Db 702 GTAMAKNSLYGCLSVHNIKYENAWYL-----WITSS 733

RESULT 7
SAKP LACSK STANDARD; PRT; 61 AA.
ID SAKP LACSK STANDARD; PRT; 61 AA.
AC P35618; Q57121;
DT 01-JUN-1994 (Rel. 29, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 13-SEP-2005 (Rel. 46, Last annotation update)
DE Bacteriocin sakacin P precursor (Sakacin 674).
GN Name=sakP; Synonyms=sakR, sppA;
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=15599;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=LTH673;
RX MEDLINE=94236240; PubMed=8180701;
RA Tichaczek P.S., Vogel R.F., Hammes W.P.;
RT "Cloning and sequencing of sakP encoding sakacin P, the bacteriocin
RT produced by Lactobacillus sake LTH 673.";
RL Microbiology 140:361-367(1994).
RN [2]
RP NUCLEOTIDE SEQUENCE, PROTEIN SEQUENCE OF 19-61, AND CHARACTERIZATION.
RC STRAIN=Lb674;
RX MEDLINE=94186010; PubMed=8138128; DOI=10.1016/0378-1097(94)90005-1;
RA Holck A.L., Axelsson L., Huehne K., Kroeckel L.;
RT "Purification and cloning of sakacin 674, a bacteriocin from
RT Lactobacillus sake Lb674.";
```

```

RL FEWS Microbiol. Lett. 115:143-150(1994).
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=Lb674;
RX MEDLINE=96262715; PubMed=8704983;
RA Huehne K., Axelsson L., Holck A., Kroeckel L.;
RT "Analysis of the sakacin P gene cluster from Lactobacillus sake Lb674
RL and its expression in sakacin-negative lb. sake strains.";
RL Microbiology 142:1437-1448(1996).
RN [4]
RP PROTEIN SEQUENCE OF 19-59, AND CHARACTERIZATION.
RC STRAIN=LTH673;
RA Tichacek P.S., Nissen-Meyer J., Nes I.P., Vogel R.F., Hammes W.P.;
RT "Characterization of the bacteriocins curvacin A from Lactobacillus
RT curvatus LTH1174 and sakacin P from L. sake LTH673.";
RL Syst. Appl. Microbiol. 15:460-465(1992).
CC -1- FUNCTION: Bactericidal activity; inhibits closely related
CC Lactobacilli, Listeria monocytogenes and ivanovii, Enterococcus
CC faecalis, Carnobacterium sp and Brochothrix thermosphacta.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: Belongs to the bacteriocin class IIA/YGNGV family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X75081; CAAS2974.1; -; Genomic DNA.
DR EMBL; Z25816; CAAS1064.1; -; Genomic DNA.
DR EMBL; Z48542; CAAS8428.1; -; Genomic DNA.
DR EMBL; AF002276; AAB93970.1; -; Genomic DNA.
DR PIR; S57911; S38508.
DR PDB; 1OG7; NMR; A=19-61.
DR PDB; 1OHM; NMR; A=19-61.
DR PDB; 1OHN; NMR; A=19-61.
DR InterPro; IPR002633; Bacteriocin_II.
DR Pfam; PF01721; Bacteriocin_II; 1.
DR ProDom; PD004452; Bacteriocin_II; 1.
KW 3D-structure; Antibiotic; Antimicrobial; Bacteriocin;
KW Direct protein sequencing.
FT PROPEP 1 18
FT CHAIN 19 61 Bacteriocin sakacin P.
FT DISULFID 27 32 By similarity.
SQ SEQUENCE 61 AA; 6385 MW; B7BF14DCDD28A73D CRC64;

Query Match 21.1%; Score 64; DB 1; Length 61;
Best Local Similarity 36.7%; Pred. No. 1.8;
Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;

Qy 3 YYGTNGVHCTKNSL---WGKVRLENKMYDQNTTYMGRL-QDILGWATG 47
||| ||||| | | | | | | | | | | | | | | | | | | | | | | |
Db 20 YYG-NGVHCGRHSCTVDWG-----TAIGNIGNNAAANWATG 54

RESULT 8
Q7B5F5 LACSK PRELIMINARY; PRT; 61 AA.
AC Q7B5F5
DT 10-MAY-2005 (TrEMBLrel. 30, Created)
DT 10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DE Sakacin P (SppA).
GN Name=sppA;
OS Lactobacillus sakei.
OC Bacteria; Firmicutes; Lactobacillales; Lactobacillaceae;
OC Lactobacillus.
OX NCBI_TaxID=1599;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MI401;
RX MEDLINE=94012314; PubMed=8407671;
RA Larsen A.G., Vogensen F.K., Josephsen J.;

```

```

RT "Antimicrobial activity of lactic acid bacteria isolated from sour
RT doughs: purification and characterization of bavaricin A, a
RL bacteriocin produced by Lactobacillus bavaricus MI401.";
RL J. Appl. Bacteriol. 75:113-122(1993).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=MI401;
RA Johansen A.H., Vogensen F.K.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=1151;
RA Urso R., Rantsiou K., Antoni C., Comi G., Luca C.;
RT "Sequencing and expression analysis of the sakacin P bacteriocin
RT produced by a Lactobacillus sakei strain isolated from naturally
RT fermented sausages.";
RL Submitted (JAN-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF526262; AM88858.1; -; Genomic DNA.
DR EMBL; AY875983; AAW79057.1; -; Genomic DNA.
SQ SEQUENCE 61 AA; 6385 MW; B7BF14DCDD28A73D CRC64;

Query Match 21.1%; Score 64; DB 2; Length 61;
Best Local Similarity 36.7%; Pred. No. 1.8;
Matches 18; Conservative 4; Mismatches 9; Indels 18; Gaps 4;

Qy 3 YYGTNGVHCTKNSL---WGKVRLENKMYDQNTTYMGRL-QDILGWATG 47
||| ||||| | | | | | | | | | | | | | | | | | | | | | | |
Db 20 YYG-NGVHCGRHSCTVDWG-----TAIGNIGNNAAANWATG 54

RESULT 9
Q51P75 MAGGR PRELIMINARY; PRT; 714 AA.
AC Q51P75
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein.
GN ORFNames=MG0601.4;
OS Magnaporthe grisea 70-15.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetes incertae sedis; Magnaporthaceae; Magnaporthe.
OX NCBI_TaxID=242507;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=70-15;
RA Birren B., Nusbaum C., Abebe A., Abouelleil A., Adekoya E.,
RA Ait-zahra M., Allen N., Allen T., An P., Anderson M., Anderson S.,
RA Arachchi H., Ambruster J., Bachantsang P., Baldwin J., Barry A.,
RA Bayul T., Blitshteyn B., Bloom I., Blye J., Boguslavskiy L.,
RA Borowsky M., Boukhgalter B., Brunache A., Butler J., Calixte N.,
RA Calvo S., Camarata J., Campo K., Chang J., Cheshatsang Y., Citroen M.,
RA Collamore A., Considine T., Cook A., Cooke P., Corum B., Cuomo C.,
RA David R., Dawoe T., Degray S., Dodge S., Dooley K., Dorje P.,
RA Dorjee K., Dorris L., Dufey N., Dupes A., Elkins T., Engela R.,
RA Erickson J., Farina A., Faro S., Ferreira P., Fischer H.,
RA Fitzgerald M., Foley K., Gage D., Galagan J., Gearin G., Gnerre S.,
RA Gnrke A., Goyette A., Graham J., Grandbois E., Gyaltsen K., Hafez N.,
RA Hagopian D., Hagos B., Hall J., Hatcher B., Heller A., Higgins H.,
RA Honan T., Horn A., Houde N., Hughes L., Hulme W., Huaby E., Iliev I.,
RA Jaffe D., Jones C., Kamal M., Katat A., Kamyselis M., Karlsson E.,
RA Kells C., Kieu A., Kisher P., Kodira C., Kulbokas E., Labutti K.,
RA Lama D., Landers T., Leger J., Levine S., Lewis D., Lewis T.,
RA Lindblad-toh K., Liu X., Lokyitsang T., Lokyitsang Y., Lucien O.,
RA Lui A., Ma L.J., Mabbitt R., Macdonald J., Maclean C., Major J.,
RA Manning J., Marabella R., Maru K., Matthews C., Mauceli E.,
RA McCarthy M., McDonough S., Mcghee T., Meldrum J., Maneva L.,
RA Mesirov J., Mihalev A., Mihova T., Mikkelsen T., Mlenga V., Moru K.,
RA Moraes J., Mulrain L., Munson G., Naylor J., Nenes C., Nguyen C.,
RA Nguyen N., Nguyen T., Nicol R., Nielsen C., Nizzari M., Norbu C.,
RA Norbu N., O'donnell P., Okoawo O., O'leary S., Omotosho B.,
RA O'Neill K., Osman S., Parker S., Perrin D., Phunkhang P., Pignani B.,
RA Purcell S., Rachupka T., Ramasamy U., Rameau R., Ray V., Raymond C.,

```

RA Retta R., Richardson S., Rise C., Rodriguez J., Rogers J., Rogov P.,  
 RA Rutnan M., Schupbach R., Seaman C., Settupalli S., Sharpe T.,  
 RA Sheridan J., Sherpa N., Shi J., Smirnov S., Smith C., Sougnez C.,  
 RA Spencer B., Stalker J., Stange-thomann N., Stavropoulos S.,  
 RA Stetson K., Stone C., Stone S., Stubbs M., Talamas J., Tchuinga P.,  
 RA Tenzing P., Tesfaye S., Theodore J., Thoulutsang Y., Topham K.,  
 RA Towey S., Tsamla T., Tesmo N., Vallee D., Vassiliev H.,  
 RA Venkataraman V., Vinson J., Vo A., Wade C., Wang S., Wangchuk T.,  
 RA Wangdi T., Whitaker C., Wilkinson J., Wu Y., Wyman D., Yadvav S.,  
 RA Yang S., Yang X., Yeager S., Yee E., Young G., Zainoun J., Zembeck L.,  
 RA Zimmer A., Zody M., Lander E.;  
 RT "The genome sequence of Magnaporthe grisea";  
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=70-15;  
 RA Dean R., Mitchell T., Brown D., Pan H., Thon M.;  
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=70-15;  
 RA Zhu H., Blackmon B.;  
 RL Submitted (OCT-2003) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an  
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is  
 CC preliminary data.  
 DR EMBL; AACU01001430; EAA48943.1; -; Genomic\_DNA.  
 KW Hypothetical protein.  
 SQ SEQUENCE 714 AA; 81698 MW; FE74ED11133DCAF3 CRC64;  
 Query Match 21.1%; Score 64; DB 2; Length 714;  
 Best Local Similarity 34.2%; Pred. No. 29;  
 Matches 13; Conservative 7; Mismatches 18; Indels 0; Gaps 0;  
 QY 7 NGVHCTKSLGKVKRLKMKYQNTTYMGRLODILLGW 44  
 DB 626 NNVHCTLSNDGTLFKSSLSDHDFYQMFVGRSDTTIHGW 663  
 RESULT 10  
 QYK80 ANASP  
 ID QYK80\_ANASP PRELIMINARY; PRT; 590 AA.  
 AC QYK80;  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)  
 DE All8037 protein.  
 GN OrderedLocusNames=all8037;  
 OS Anabaena sp. (strain PCC 7120).  
 OG Plasmid pCC7120gamma.  
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.  
 OX NCBI\_TaxID=103690;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RX MEDLINE=21595285; PubMed=11759840;  
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,  
 RA Watanabe A., Iriiguchi M., Ishikawa A., Kawashima K., Kimura T.,  
 RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,  
 RA Nakazaki N., Shimpo S., Sugimoto M., Takazawa M., Yamada M.,  
 RA Yasuda M., Tabata S.;  
 RA "Complete genomic sequence of the filamentous nitrogen-fixing  
 RT cyanobacterium Anabaena sp. strain PCC 7120.";  
 RL DNA Res. 8:205-213(2001).  
 DR EMBL; AP003603; BAB77367.1; -; Genomic\_DNA.  
 DR PIR; AF2555; AF2555.  
 DR GO; GO:0016020; C:membrane; IEA.  
 DR GO; GO:0009291; P:unidirectional conjugation; IEA.  
 KW Complete proteome; Plasmid.  
 SQ SEQUENCE 590 AA; 64933 MW; 7ECA68DF6FE3B469 CRC64;  
 Query Match 20.7%; Score 63; DB 2; Length 590;  
 Best Local Similarity 35.3%; Pred. No. 31;  
 Matches 18; Conservative 8; Mismatches 23; Indels 0; Gaps 1;

QY 3 YYGTNGVHCTKSLGKVKRLKMKYQNTTYMGRLODILLGWATGAFGKTF 53  
 DB 88 YIGT--PRGTFQVGVGNKRITNIPEDKRNRLYLPDVQRGILVSGSGSGKTF 136  
 RESULT 11  
 QYRIM1 FUSNN  
 ID QYRIM1\_FUSNN PRELIMINARY; PRT; 1582 AA.  
 AC QYRIM1;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Fusobacterium outer membrane protein family.  
 GN OrderedLocusNames=FN1554;  
 OS Fusobacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;  
 CC Fusobacterium.  
 OX NCBI\_TaxID=76856;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ATCC 25586;  
 RX MEDLINE=21886394; PubMed=11889109;  
 RA DOI=10.1128/JB.184.7.2005-2018.2002;  
 RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,  
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,  
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,  
 RA Fonstein M., Kyrpides N.C., Overbeek R.;  
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium  
 RT nucleatum strain ATCC 25586.";  
 RL J. Bacteriol. 184:2005-2018(2002).  
 DR EMBL; AE009951; AAL93680.1; -; Genomic\_DNA.  
 DR InterPro; IPR001014; AutoTranspheta.  
 DR Pfam; PF03797; Autotransporter; 1.  
 DR PROSITE; PS00050; RIBOSOMAL\_L23; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 1582 AA; 167889 MW; 666C38783A00EB4F CRC64;  
 Query Match 20.7%; Score 63; DB 2; Length 1582;  
 Best Local Similarity 35.2%; Pred. No. 94;  
 Matches 19; Conservative 6; Mismatches 21; Indels 8; Gaps 3;  
 QY 4 YGTNGVHCTKSLGKVKRLKMKYQNTTYMGRLODILL---GWATGAFGKTF 53  
 DB 1310 FGTNGEYKTDTA--GVIDYKNHAY--GVAYVHNEDEIKLGRGTGWTGIVHNTF 1359  
 RESULT 12  
 QYRHH7 FUSNN  
 ID QYRHH7\_FUSNN PRELIMINARY; PRT; 1630 AA.  
 AC QYRHH7;  
 DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Fusobacterium outer membrane protein family.  
 GN OrderedLocusNames=FN2047;  
 OS Fusobacterium nucleatum (subsp. nucleatum).  
 OC Bacteria; Fusobacteria; Fusobacteriales; Fusobacteriaceae;  
 CC Fusobacterium.  
 OX NCBI\_TaxID=76856;  
 RN [1]  
 RP NUCLEOTIDE SEQUENCE.  
 RC STRAIN=ATCC 25586;  
 RX MEDLINE=21886394; PubMed=11889109;  
 RA DOI=10.1128/JB.184.7.2005-2018.2002;  
 RA Kapral V., Anderson I., Ivanova N., Reznik G., Los T., Lykidis A.,  
 RA Bhattacharyya A., Bartman A., Gardner W., Grechkin G., Zhu L.,  
 RA Vasileva O., Chu L., Kogan Y., Chaga O., Goltsman E., Bernal A.,  
 RA Larsen N., D'Souza M., Walunas T., Pusch G., Haselkorn R.,  
 RA Fonstein M., Kyrpides N.C., Overbeek R.;  
 RT "Genome sequence and analysis of the oral bacterium Fusobacterium

```

RT nucleatum strain ATCC 25586.";
RL J. Bacteriol. 184:2005-2018(2002).
DR EMBL; AE009951; AAL94131.1; -; Genomic_DNA.
DR InterPro; IPR005546; Auto_transp_beta.
DR Pfam; PF03797; Autotransporter; 1.
KW Complete proteome.
SQ SEQUENCE 1630 AA; 172743 MW; 38E9CE42F037B75A CRC64;

Query Match 20.7%; Score 63; DB 2; Length 1630;
Best Local Similarity 35.2%; Pred. No. 97;
Matches 19; Conservative 6; Mismatches 21; Indels 8; Gaps 3;

QY 4 YGTNGVHCTKSLWGVRLKNKYDQNTYMGRLQDILL----GWATGAFGKTF 53
Db 1358 FCTNGEYKTDTA--GVIDYKHAY--GVAYVHENEIKLGRIGWYTGIVHNTF 1407

RESULT 13
GLNA_METVO STANDARD; PRT; 446 AA.
AC P21154;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Glutamine synthetase (EC 6.3.1.2) (Glutamate--ammonia ligase) (GS).
GN Name=gluA;
OS Methanococcus voltae.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanococcaceae; Methanococcus.
OX NCBI_TaxID=2188;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=DSM 1537 / PS;
RX MEDLINE=90139872; PubMed=2575777; DOI=10.1016/0923-2508(89)90012-0;
RA Posset O., Sibold L., Aubert J.-P.;
RT "Nucleotide sequence and expression of the glutamine synthetase structural gene, glnA, of the archaeobacterium Methanococcus voltae.";
RL Res. Microbiol. 140:355-371(1989).
CC -!- CATALYTIC ACTIVITY: ATP + L-glutamate + NH(3) = ADP + phosphate + L-glutamine.
CC -!- ENZYME REGULATION: The activity of this enzyme is controlled by adenylation under conditions of abundant glutamine. The fully adenylated enzyme complex is inactive (By similarity).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- SIMILARITY: Belongs to the glutamine synthetase family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use as long as its content is in no way modified and this statement is not removed.
CC -----
DR EMBL; X53509; CAA37585.1; -; Genomic_DNA.
DR PIR; A43995; A43995.
DR HSP; P06201; ILGR.
DR InterPro; IPR008147; Gln_synt_beta.
DR InterPro; IPR008146; Gln_synt_C.
DR InterPro; IPR004809; GluA.
DR Pfam; PF00120; Gln-synt_C; 1.
DR Pfam; PF03951; Gln-synt_N; 1.
DR ProDom; PD001057; Gln_synt_C; 1.
DR TIGRFAMs; TIGR00653; GluA; 1.
DR PROSITE; PS00180; GLNA_1; 1.
DR PROSITE; PS00181; GLNA_ATP; 1.
KW Ligase.
FT BINDING 396 396 AMP (covalent) (By similarity).
SQ SEQUENCE 446 AA; 50199 MW; 95449E4DE8542690 CRC64;

Query Match 20.6%; Score 62.5; DB 1; Length 446;
Best Local Similarity 33.9%; Pred. No. 26;
Matches 19; Conservative 9; Mismatches 17; Indels 11; Gaps 5;

QY 1 KTYGYTN--GVHCTKNSLW--GKVLKNKYDQNTYMGRLQDILLGWATGAFGKT 52

```

```

Db 237 KPFFGMNGSGMHCNQ-SIWLDGKPSF-----YDENNAH--QLSDICLSYIGILEHT 285

RESULT 14
Q8VM63_BACTA PRELIMINARY; PRT; 558 AA.
AC Q8VM63;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Cry39ORF2 protein.
GN Name=Cry3orf2;
OS Bacillus thuringiensis (subsp. aizawai).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus;
OC Bacillus cereus group.
OX NCBI_TaxID=1433;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ito T., Sahara K., Bando H., Asano S.;
RT "Cloning and Expression of Novel Crystal Protein Genes cry39A and cry39f2 from Bacillus thuringiensis subsp. aizawai Bn1-14 Encoding Mosquitocidal Proteins.";
RL J. Insect Biotechnol. Sericulture 71:123-128(2002).
DR EMBL; AB074413; BAB72017.1; -; Genomic_DNA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
SQ SEQUENCE 558 AA; 63360 MW; 26F6D5A013834207 CRC64;

Query Match 20.6%; Score 62.5; DB 2; Length 558;
Best Local Similarity 27.7%; Pred. No. 34;
Matches 13; Conservative 8; Mismatches 17; Indels 9; Gaps 2;

QY 9 VHCTKNS-----LMGKVR-LKNMKYDQNTYMGRLQDILLGWAT 46
Db 26 IECMSNEHSSKEENMLMDEVKQAKQLSWRNLLYNGDFEDVSNWTKT 72

RESULT 15
Q9PZ03_GVXN PRELIMINARY; PRT; 469 AA.
AC Q9PZ03;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF40.
GN Name=ORF40;
OS Xestia c-nigrum granulosis virus (XnGV) (Xestia c-nigrum granulovirus).
OC Viruses; dsDNA viruses, no RNA stage; Baculoviridae; Granulovirus.
OX NCBI_TaxID=51677;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Hayakawa T., Ko R., Okano K., Seong S.-I., Goto C., Maeda S.;
RT "Sequence analysis of the Xestia c-nigrum granulovirus genome.";
RL Virology 262:277-297(1999).
DR EMBL; AF162221; AAF05154.1; -; Genomic_DNA.
DR HSP; P03956; 1CGL.
DR MEROPS; M10.032; -.
DR GO; GO:0005578; C:extracellular matrix (sensu Metazoa); IEA.
DR GO; GO:0004222; F:metalloendopeptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR000585; Hemopeptidase M.
DR InterPro; IPR001818; Pept_M10A_M12B.
DR InterPro; IPR006025; Pept_M_Zn_BS.
DR Pfam; PF00413; Peptidase_M10; 1.
DR PRINTS; PR00138; MATRXIN.
DR SMART; SM00235; ZmMC; 1.
DR PROSITE; PS00142; ZINC_PROTEASE; UNKNOWN 1.
SQ SEQUENCE 469 AA; 53808 MW; 1BF4EB5C210E241 CRC64;

Query Match 20.2%; Score 61.5; DB 2; Length 469;

```

Best Local Similarity	35.9%;	Pred. No. 38;
Matches	14; Conservative	9; Mismatches
		7; Indels
		9; Gaps
		2;

Qy	5	GTNGVHCTKNSLWGVRLKNNKYDQNTTYM--GRLQDIL	41
		: :   :   :   :   :   :   :   :   :	
Db	335	GDEIATRNLLW-----YEHKNSLTMNVGRVQDVL	366

Search completed: May 2, 2006, 17:30:32  
Job time : 237 secs

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:30:50 ; Search time 46 Seconds  
(without alignments)  
97.054 Million cell updates/sec

Title: US-10-644-927-1  
Perfect score: 304  
Sequence: 1 KTYGNGVHCTKNSLWGV.....GRQDILLGWATGAFGKTEH 54

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

```

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                  Maximum Match 10%
                  Listing first 45

```

```
Database : Issued Patents_AA:*
1: /cgn2_6/pdata1/1/aaa/5_COMB pep.*
2: /cgn2_6/pdata1/1/aaa/6_COMB pep.*
3: /cgn2_6/pdata1/1/aaa/H_COMB pep.*
4: /cgn2_6/pdata1/1/aaa/PCTUS_COMB pep.*
5: /cgn2_6/pdata1/1/aaa/RE_COMB pep.*
6: /cgn2_6/pdata1/1/aaa/backfiles1 pep.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	69.5	22.9	420	2	US-09-328-352-5261	Sequence 5261, Appl
2	64	21.1	61	2	US-09-068-507C-5	Sequence 5, Appl1
3	62	20.4	85	2	US-09-270-767-44549	Sequence 44549, A
4	60	19.7	36	2	US-08-924-629C-69	Sequence 69, Appl
5	60	19.7	37	2	US-09-030-619-209	Sequence 209, Appl
6	60	19.7	37	2	US-09-444-281-90	Sequence 90, Appl
7	60	19.7	61	2	US-08-924-629C-68	Sequence 68, Appl
8	60	19.7	300	2	US-09-982-616-9	Sequence 9, Appl1
9	59.5	19.6	577	2	US-09-248-796A-18807	Sequence 18807, A
10	59	19.4	62	2	US-08-924-629C-71	Sequence 71, Appl
11	56.5	18.6	242	2	US-09-543-681A-4928	Sequence 4928, Ap
12	56	18.4	66	2	US-09-107-532A-4680	Sequence 4680, Ap
13	56	18.4	909	2	US-09-982-616-11	Sequence 11, Appl
14	56	18.4	10182	2	US-09-134-001C-3159	Sequence 3159, Ap
15	55.5	18.3	243	2	US-09-216-295-13	Sequence 13, Appl
16	55.5	18.3	244	2	US-09-632-570-13	Sequence 13, Appl
17	55.5	18.3	244	2	US-09-632-575-43	Sequence 43, Appl
18	55	18.1	41	2	US-08-924-629C-70	Sequence 70, Appl
19	55	18.1	177	2	US-08-248-796A-18521	Sequence 18521, A
20	54.5	17.9	494	2	US-08-378-313-23	Sequence 23, Appl
21	54.5	17.9	494	2	US-08-378-313-29	Sequence 29, Appl
22	54.5	17.9	641	2	US-09-071-035-456	Sequence 456, App
23	54.5	17.9	641	2	US-10-206-576-456	Sequence 456, App
24	54.5	17.9	1313	2	US-09-071-035-450	Sequence 450, App
25	54.5	17.9	1313	2	US-09-071-035-454	Sequence 454, App
26	54.5	17.9	1313	2	US-10-206-576-450	Sequence 450, App
27	54.5	17.9	1313	2	US-10-206-576-454	Sequence 454, App

## ALIGNMENTS

## RESULT 1

```

US-09-328-352-5261
; Sequence 5261, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSIS
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 5261
; LENGTH: 420
; TYPE: PRN
; ORGANISM: Acinetobacter baumannii
US-09-328-352-5261

```

Query Match	22.9%;	Score 69.5;	DB 2;	Length 420;
Best Local Similarity	38.5%;	Pred. NO. 0.47;		
Matched	20	8. Mismatches	13:	Indels 11:
Conservative				

```
Qy      1 KTYGTN---GVHCTKNSLWGVRLKNMKY--DQNTTVMGRLQDI---LLGW 44
      :: ||| | | | | | | | | | | | | | | | | | | | | |
ph     361 FSVVVTNPKSGCHSNKANIWK---NNTPYSGDRDPTVRLRLDITSPILGW 409
```

RESIT.T 2

```

RES001.2
US-09-068-507C-5
; Sequence 5, Application US/09068507C
; Patent No. 6790951
; GENERAL INFORMATION:
; APPLICANT: ELSINK, VINCENT et al.
; TITLE OF INVENTION: EXPRESSION SYSTEM
; TITLE OF INVENTION: HETEROLOGOUS N
; FILE REFERENCE: 1380-0122P
; CURRENT APPLICATION NUMBER: US/09/0
; CURRENT FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Lactobacillus sake
US-09-068-507C-5

```

Query Match	21.1%;	Score 64;	DB 2;	Length 61;
Best Local Similarity	36.7%;	Pred. No. 0.23;		
Matches 18:	Conservative	4;	Mismatches	9;
			Indels	18;
			Gaps	4;

Sequence 7485, Appl  
Sequence 3, Appl  
Sequence 5289, Appl  
Sequence 4, Appl  
Sequence 15, Appl  
Sequence 1, Appl  
Sequence 12, Appl  
Sequence 24, Appl  
Sequence 24, Appl  
Sequence 54, Appl  
Sequence 2, Appl  
Sequence 6424, Appl  
Sequence 17748, A  
Sequence 12, Appl  
Sequence 6, Appl  
Sequence 2, Appl  
Sequence 2633, Appl  
Sequence 37353, A  
Sequence 37353, A

**Qy** 3 YYGNGVHCTKNSL---WGKVRLLKNMKYDQNTTYMGRLL-QDILLGWATG 47  
||| ||||| :| | | |  
**Pb** 20 YYG-NGVHGHSCTVDWG-----TAIGNGNNAANWATG 54  
||| ||||| :| | | |

```

RESULT 3
US-09-270-767-44549
; Sequence 44549, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 44549
; LENGTH: 85
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-44549

```

```

Query Match      20.4%; Score 62; DB 2; Length 85;
Best Local Similarity 38.2%; Pred. No. 0.66;
Matches 13; Conservative 4; Mismatches 17; Indels 0; Gaps 0;

Qy  7  NGVCTNSLWGGKVRLLKNMKYDQNTTYMGRLQDI  40
      | | | | | | | | | | | | | | | | | |
Db  46  NACHCNKCSKWSYXKVVIRKYFOOGSSBEGVLNDI  79

```

RESULT 4  
US-08-924-629C-69  
; Sequence 69, Application US/08924629C  
; Patent No. 6403082  
; GENERAL INFORMATION:  
; APPLICANT: Stiles, Michael E.  
; APPLICANT: Vederas, John C.  
; APPLICANT: van Belkum, Marius J.  
; APPLICANT: Worobo, Randy W.  
; APPLICANT: Worobo, Rodney J.  
; APPLICANT: Greer, G. Gordon  
; APPLICANT: McMullen, Lynn M.  
; APPLICANT: Leisner, Jorgen J.  
; APPLICANT: Poon, Alesion  
; APPLICANT: Franz, Charles M.A.P.  
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method  
; FILE REFERENCE: 660.0005US  
; CURRENT APPLICATION NUMBER: US/08/924,629C  
; CURRENT FILING DATE: 1997-09-05  
; PRIOR APPLICATION NUMBER: US 60/026,257  
; PRIOR FILING DATE: 1996-09-05  
; NUMBER OF SEQ ID NOS: 80  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 69  
; LENGTH: 36  
; TYPE: prt  
; ORGANISM: Mesenteriocin Y105  
; US-08-924-629C-69

```

Query Match      19.7%; Score 60; DB 2; Length 36;
Best Local Similarity 48.4%; Pred. No. 0.42;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

Qy  3  YYGTNGVHCTKNSL---WGKV-----RLKN  24
      ||| ||| ||| |||
Db   2  YYG-NGVHCTKGGCGSVNKGGAASAGIIRLAN  31
      ||| ||| ||| |||

```

## RESULT 5

```

US-09-030-619-209
; Sequence 209, Application US/09030619B
; Patent No. 6503881
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Ertle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406
; CURRENT APPLICATION NUMBER: US/09/030,619B
; CURRENT FILING DATE: 1998-02-25
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidium
; US-09-030-619-209

```

```

Query Match      19.7%; Score 60; DB 2; Length 37;
Best Local Similarity 48.4%; Pred. No. 0.44;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

Qy      3  YYGTNGVHCTKNSL---WGKV-----RLKN 24
Db      2  YYG-NGVHCTKSGCSVNWGEAFSAGVHRLAN 31

```

```

RESULT 6
US-09-444-281-90
; Sequence 90, Application US/09444281
; Patent No. 6946261
; GENERAL INFORMATION:
; APPLICANT: Burian, Jan
; APPLICANT: Bartfeld, Daniel
; TITLE OF INVENTION: EFFICIENT METHODS FOR PRODUCING
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES IN HOST CELLS
; FILE REFERENCE: 660081.411
; CURRENT APPLICATION NUMBER: US/09/444.281
; CURRENT FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 90
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidium
US-09-444-281-90

```

Query Match 19.7%; Score 60; DB 2; Length 37;  
Best Local Similarity 48.4%; Pred. No. 0.44;  
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

Qy 3 YYGTNGVHCTKNSL---WGKV-----RLKN 24  
||| |||  
Db 2 YYG-NGVHCTKSGCSVNWGEAFSAGVHRLAN 31

RESULT 7  
US-08-924-629C-68  
; Sequence 68: Application US/08924629C  
; Patent No. 6403082  
; GENERAL INFORMATION:  
; APPLICANT: Stiles, Michael E.  
; APPLICANT: Vederas, John C.  
; APPLICANT: van Belkum, Marius J.  
; APPLICANT: Worobo, Randy W.  
; APPLICANT: Worobo, Rodney J.  
; APPLICANT: Greer, G. Gordon



```
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aislon
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924.629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Leucococin A
US-08-924-629C-68

Query Match 19.7%; Score 60; DB 2; Length 61;
Best Local Similarity 48.4%; Pred. No. 0.82;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

Qy 3 YGTNGVHCTKNSL-----WGKV-----RLKN 24
Db 26 YYG-NGVHCTKSGCSVNWGEAESAGVHRLAN 55

RESULT 8
US-09-982-616-9
; Sequence 9, Application US/09982616
; Patent No. 6955884
; GENERAL INFORMATION:
; APPLICANT: Frances E. Lund
; APPLICANT: Troy D. Randall
; APPLICANT: Santiago Partida-Sanchez
; TITLE OF INVENTION: CD38 MODULATED CHEMOTAXIS
; FILE REFERENCE: AP33438 068443.0106
; CURRENT APPLICATION NUMBER: US/09/982,616
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/241,065
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-982-616-9

Query Match 19.7%; Score 60; DB 2; Length 300;
Best Local Similarity 35.0%; Pred. No. 6.1;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy 5 GTNGVHCTKNSLWGKVRLLKNMKYDQNTTYMGRLLQDILLGW 44
Db 113 GTQTVPCKILLWSRIKDLAHQFTQVRDMFTLEDTLGCT 152

RESULT 9
US-09-248-796A-18807
; Sequence 18807, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
```

```
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 18807
; LENGTH: 577
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-18807

Query Match 19.6%; Score 59.5; DB 2; Length 577;
Best Local Similarity 30.6%; Pred. No. 16;
Matches 15; Conservative 7; Mismatches 12; Indels 15; Gaps 2;

Qy 6 TNGVHCTKNSLWGKVRLLKNMKYDQNTTYMGRLLQDILLGWATGAFGKTFH 54
Db 261 TTGVQCDESTLWVKL-IPNLKHLNN-----QTSAIGKFIH 294

RESULT 10
US-08-924-629C-71
; Sequence 71, Application US/08924629C
; Patent No. 6403082
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Aislon
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. 6403082elBacteriocins, Transport and Vector System and Method
; FILE REFERENCE: 660.0005US
; CURRENT APPLICATION NUMBER: US/08/924,629C
; CURRENT FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 71
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Pediocin PA1
US-08-924-629C-71

Query Match 19.4%; Score 59; DB 2; Length 62;
Best Local Similarity 37.5%; Pred. No. 1.1;
Matches 18; Conservative 5; Mismatches 9; Indels 16; Gaps 4;

Qy 3 YGTNGVHCTKNSL---WGKVRLLKNMKYDQNTTYMGRLLQDILLGWATG 47
Db 20 YYG-NGVTCGRHSCSVDMGKA-----TTTCI--INNGNMAWATG 54

RESULT 11
US-09-543-681A-4928
; Sequence 4928, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABIL
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4928
; LENGTH: 242
; TYPE: PRT
; ORGANISM: Proteus mirabilis
; FEATURE:
```

```
; NAME/KEY: UNSURE
; LOCATION: (18)
; OTHER INFORMATION: Identity of amino acid at the above locations are unknown.
US-09-543-681A-4928

Query Match      18.6%   Score 56.5; DB 2; Length 242;
Best Local Similarity 28.1%; Pred.No.14;
Matches 18; Conservative 8; Mismatches 21; Indels 17; Gaps 3;

QY    4 YTGNGVHCTKNSLWGVRLKN-----MKYDQNTTYMGRLODLILGNATGAF 49
     ||| :||: | || 
Db    106 YKTSSTDSAKISVFGLVNAKNTEIGGYIIDVIDIKVDNKRAYLGSRVADKI--W-TGSS 162
       ||| :||: | || 

QY    50 GKTF 53
           |
Db    163 ATIF 166

RESULT 12
US-09-107-532A-4680
; Sequence 4680, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
          ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
          ADDRESSEE: GENOME THERAPEUTICS CORPORATION
          STREET: 100 Beaver Street
          CITY: Waltham
          STATE: Massachusetts
          COUNTRY: USA
          ZIP: 02354
; COMPUTER READABLE FORM:
          MEDIUM TYPE: CD-ROM ISO9660
          COMPUTER: PC
          OPERATING SYSTEM: <Unknown>
          SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
          FILING DATE: July 2, 1997
          PRIORITY DATE: 30-Jun-1998
          PRIOR APPLICATION DATA:
          APPLICATION NUMBER: 60/085,598
          FILING DATE: 14 May 1998
          APPLICATION NUMBER: 60/051571
          FILING DATE: July 2, 1997
          ATTORNEY/AGENT INFORMATION:
          NAME: Ariniello, Pamela Deneke
          REGISTRATION NUMBER: 40,489
          REFERENCE/DOCKET NUMBER: GTC-012
          TELECOMMUNICATION INFORMATION:
          TELEPHONE: (781) 893-5007
          TELEFAX: (781) 893-8277
          INFORMATION FOR SEQ ID NO: 4680:
            SEQUENCE CHARACTERISTICS:
              LENGTH: 66 amino acids
              TYPE: amino acid
              TOPOLOGY: linear
              MOLECULE TYPE: protein
              HYPOTHETICAL: YES
              ORIGINAL SOURCE:
                ORGANISM: Enterococcus faecium
              FEATURE:
                NAME/KEY: misc feature
                LOCATION: (B) LOCATION 1...66
                SEQUENCE DESCRIPTION: SEQ ID NO: 4680:
US-09-107-532A-4680
```

```

Qy      3 YGTNGVHCTKNSL---WGK 19
      ||||| :|||:|||||
Db      26 YYG-NGVYTKNKCTVDWAK 44

RESULT 13
US-09-982-616-11
; Sequence 11, Application US/09982616
; Patent No. 6955884
; GENERAL INFORMATION:
; APPLICANT: Frances E. Lund
; APPLICANT: Troy D. Randall
; APPLICANT: Santiago Partida-Sanchez
; TITLE OF INVENTION: CD38 MODULATED CHEMOTAXIS
; FILE REFERENCE: AP33438 068443.0106
; CURRENT APPLICATION NUMBER: US/09/982.616
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/241,065
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 909
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Reverse translation of SM38
US-09-982-616-11

Query Match          18.4%; Score 56; DB 2; Length 909;
Best Local Similarity 32.7%; Pred. No. 85;
Matches 16; Conservative 2; Mismatches 15; Indels 16; Gaps 2;

Qy      4 YGTNGVHCTKNSLWGVRLKNMKYDQNTTYMGRLQDILLG--WATGAFG 50
      ||||| :|||:|||||
Db      477 YGNGTINCA-----YGCNTTYGGCARWSNGCWSNGCG 511

RESULT 14
US-09-134-001C-3159
; Sequence 3159, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GPC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; CURRENT FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 3159
; LENGTH: 10182
; TYPE: PRT
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-3159

Query Match          18.4%; Score 56; DB 2; Length 10182;
Best Local Similarity 46.7%; Pred. No. 1.8e+03;
Matches 14; Conservative 2; Mismatches 14; Indels 0; Gaps 0;

Qy      6 TNGVHCTKNSLWGVRLKNMKYDQNTTYMG 35
      ||||| :|||:|||||
Db      4753 TORVNTTKNDLNGNDKLAEAKRDANTTIDG 4782

RESULT 15
US-09-216-295-13
; Sequence 13, Application US/09216295
; Patent No. 6268328
```

; GENERAL INFORMATION:  
; APPLICANT: Mitchinson, Colin  
; APPLICANT: Wendt, Dan J.  
; TITLE OF INVENTION: No. 6268328el Variant EGIII-Like Cellulase Compositions  
; FILE REFERENCE: GC555  
; CURRENT APPLICATION NUMBER: US/09/216,295  
; CURRENT FILING DATE: 1998-12-18  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 13  
; LENGTH: 243  
; TYPE: PRT  
; ORGANISM: Fusarium javanicum (1)  
US-09-216-295-13

Query Match 18.3%; Score 55.5; DB 2; Length 243;  
Best Local Similarity 31.4%; Pred. No. 19;  
Matches 16; Conservative 8; Mismatches 20; Indels 7; Gaps 2;  
Qy 6 TNGVHCTKNSLWGVRLKMKYDONTTMYGRLQDILLGWA-----TGAFGK 51  
Db 34 TAGAYTIYNLWGVKDAES--GEQCTTNSGEQSDGSIWVSWTGGQQG 82

Search completed: May 2, 2006, 17:32:07  
Job time : 47 secs

**THIS PAGE LEFT BLANK**

Wed May 3 10:09:57 2006

us-10-644-927-1.rapbm

Page 1

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:42:01 ; Search time 164 Seconds  
(without alignments)  
137.578 Million cell updates/sec

Title: US-10-644-927-1

Perfect score: 304  
Sequence: 1 KTYGTNGVHCTKNSLWKV.....GRLQDILLGWATGAFGKTFH 54

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA Main:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pap.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pap.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pap.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pap.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pap.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US11\_PUBCOMB.pap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	304	100.0	54	5	US-10-644-927-1
2	62.5	20.6	558	5	US-10-781-979-27
3	60	19.7	36	3	US-09-883-343A-69
4	60	19.7	37	3	US-09-030-619-209
5	60	19.7	37	3	US-09-912-609-97
6	60	19.7	37	4	US-10-277-232-209
7	60	19.7	37	4	US-10-277-233-209
8	60	19.7	61	3	US-09-883-343A-68
9	60	19.7	300	3	US-09-982-616-9
10	60	19.7	300	3	US-09-773-307B-2
11	60	19.7	300	4	US-10-295-027-1348
12	60	19.7	300	4	US-10-454-238-2
13	60	19.7	300	4	US-10-322-696-162
14	60	19.7	300	5	US-10-753-267-34
15	60	19.7	688	5	US-10-450-763-37242
16	59	19.4	62	3	US-09-883-343A-71
17	59	19.4	399	4	US-10-369-493-6141
18	59	19.4	753	5	US-10-983-198-42
19	58.5	19.2	95	4	US-10-424-599-280304
20	58	19.1	511	4	US-10-424-599-213009
21	57.5	18.9	461	4	US-10-410-432-10
22	57.5	18.9	461	4	US-10-669-174-10
23	57.5	18.9	461	5	US-10-984-956A-10
24	57.5	18.9	461	5	US-10-732-923-19385
25	57.5	18.9	643	6	US-11-097-143-40056
26	57	18.8	564	4	US-10-424-599-197752
27	57	18.8	581	4	US-10-282-122A-72532

## ALIGNMENTS

## RESULT 1

US-10-644-927-1

; Sequence 1, Application US/10644927

; Publication No. US20050153881A1

; GENERAL INFORMATION:

; APPLICANT: Stern, Norman J

; APPLICANT: Svetoch, Edward A.

; APPLICANT: Eruslanov, Boris V.

; APPLICANT: Volodina, Larisa I.

; APPLICANT: Kovalev, Yuri N.

; APPLICANT: Kudryavtseva, Tamara Y.

; APPLICANT: Pereygin, Vladimir V.

; APPLICANT: Pokhilenko, Victor D.

; APPLICANT: Levchuk, Vladimir P.

; APPLICANT: Borzenkov, Valery N.

; APPLICANT: Svetoch, Olga E.

; APPLICANT: Mitsevich, Eugeni V.

; APPLICANT: Mitsevich, Irina P.

; TITLE OF INVENTION: Bacteriocins and Novel Bacterial Strains

; FILE REFERENCE: D.N. 0135.03

; CURRENT APPLICATION NUMBER: US/10/644,927

; CURRENT FILING DATE: 2003-08-21

; NUMBER OF SEQ ID NOS: 5

; SOFTWARE: Patent in version 3.2

; SEQ ID NO 1

; LENGTH: 54

; TYPE: PRT

; ORGANISM: Lactobacillus salivarius

; US-10-644-927-1

Query Match 100.0%; Score 304; DB 5; Length 54;

Best Local Similarity 100.0%; Pred. No. 1.4e-31;

Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 KTYGTNGVHCTKNSLWKVRLKNMKYDONTTYMGRLODILLGWATGAFGKTFH 54

Db 1 KTYGTNGVHCTKNSLWKVRLKNMKYDONTTYMGRLODILLGWATGAFGKTFH 54

## RESULT 2

US-10-781-979-27

; Sequence 27, Application US/10781979

; Publication No. US20040250311A1

; GENERAL INFORMATION:

; APPLICANT: Carozzi, Nadine

; APPLICANT: Hargiss, Tracy

; APPLICANT: Koziel, Michael G.

; APPLICANT: Duck, Nicholas B.

; APPLICANT: Carr, Brian

; TITLE OF INVENTION: AXMI-008, A Delta-Endotoxin Gene and

Sequence 225655, Sequence 11, Appl  
Sequence 70580, A  
Sequence 23, Appl  
Sequence 4098, Ap  
Sequence 338816,  
Sequence 304001,  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 70314, A  
Sequence 251103,  
Sequence 60740, A  
Sequence 70, Appl  
Sequence 5582, Ap  
Sequence 203257,  
Sequence 55991, A  
Sequence 103806,  
Sequence 27507, A

28 56 18.4 636 4 US-10-424-599-225665  
29 56 18.4 909 3 US-09-982-616-11  
30 56 18.4 6641 4 US-10-282-122A-70580  
31 56 18.4 10203 4 US-10-661-809-23  
32 56 18.4 10203 4 US-10-724-972A-4098  
33 55.5 18.3 81 4 US-10-425-115-338836  
34 55.5 18.3 108 4 US-10-425-115-304001  
35 55.5 18.3 244 4 US-10-441-626-13  
36 55.5 18.3 244 4 US-10-441-626-13  
37 55.5 18.3 379 4 US-10-425-114-70314  
38 55.5 18.3 553 4 US-10-425-115-251103  
39 55.5 18.3 608 4 US-10-425-114-60740  
40 55 18.1 41 3 US-09-883-343A-70  
41 55 18.1 112 3 US-09-738-626-5582  
42 55 18.1 153 4 US-10-425-115-203257  
43 55 18.1 285 4 US-10-282-122A-55991  
44 55 18.1 2071 4 US-10-437-963-103806  
45 54.5 17.9 302 6 US-11-097-143-27507



; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406C1
; CURRENT APPLICATION NUMBER: US/10/277,232
; CURRENT FILING DATE: 2002-11-27
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidum
US-10-277-232-209

Query Match 19.7%; Score 60; DB 4; Length 37;
Best Local Similarity 48.4%; Pred. No. 3.1;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

QY 3 YGTNGVHCTKNSL---WGKV-----RLKN 24
DB 2 YYG-NGVHCTKSGCSVNWGEAFSAGVHRLAN 31

RESULT 7
US-10-277-233-209
; Sequence 209, Application US/10277233
; Publication No. US20030232750A1
; GENERAL INFORMATION:
; APPLICANT: Krieger, Timothy J.
; APPLICANT: Taylor, Robert
; APPLICANT: Erle, Douglas
; APPLICANT: Fraser, Janet R.
; APPLICANT: West, Michael H.P.
; APPLICANT: McNicol, Patricia J.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR TREATING
; TITLE OF INVENTION: INFECTIONS USING CATIONIC PEPTIDES ALONE OR IN COMBINATION
; TITLE OF INVENTION: WITH ANTIBIOTICS
; FILE REFERENCE: 660081.406C1
; CURRENT APPLICATION NUMBER: US/10/277,233
; CURRENT FILING DATE: 2002-10-18
; NUMBER OF SEQ ID NOS: 232
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 209
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidum
US-10-277-233-209

Query Match 19.7%; Score 60; DB 4; Length 37;
Best Local Similarity 48.4%; Pred. No. 3.1;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

QY 3 YGTNGVHCTKNSL---WGKV-----RLKN 24
DB 2 YYG-NGVHCTKSGCSVNWGEAFSAGVHRLAN 31

RESULT 8
US-09-883-343A-68
; Sequence 68, Application US/09883343A
; Publication No. US20030039632A1
; GENERAL INFORMATION:
; APPLICANT: Stiles, Michael E.
; APPLICANT: Vederas, John C.
; APPLICANT: van Belkum, Marius J.
; APPLICANT: Worobo, Randy W.
; APPLICANT: Worobo, Rodney J.
; APPLICANT: Greer, G. Gordon
; APPLICANT: McMullen, Lynn M.
; APPLICANT: Leisner, Jorgen J.
; APPLICANT: Poon, Alston
; APPLICANT: Franz, Charles M.A.P.
; TITLE OF INVENTION: No. US20030039632A1elBacteriocins, Transport and Vector System an
; FILE REFERENCE: 660.0005US

; CURRENT APPLICATION NUMBER: US/09/883,343A
; CURRENT FILING DATE: 2001-06-19
; PRIOR APPLICATION NUMBER: US/08/924,629
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 60/026,257
; PRIOR FILING DATE: 1996-09-05
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 68
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Leucocin A
US-09-883-343A-68

Query Match 19.7%; Score 60; DB 3; Length 61;
Best Local Similarity 48.4%; Pred. No. 5.5;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

QY 3 YGTNGVHCTKNSL---WGKV-----RLKN 24
DB 26 YYG-NGVHCTKSGCSVNWGEAFSAGVHRLAN 55

RESULT 9
US-09-982-616-9
; Sequence 9, Application US/09982616
; Publication No. US20020127646A1
; GENERAL INFORMATION:
; APPLICANT: Frances E. Lund
; APPLICANT: Troy D. Randall
; APPLICANT: Santiago Partida-Sanchez
; TITLE OF INVENTION: CD38 MODULATED CHEMOTAXIS
; FILE REFERENCE: AP33438 068443.0106
; CURRENT APPLICATION NUMBER: US/09/982,616
; CURRENT FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/241,065
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-982-616-9

Query Match 19.7%; Score 60; DB 3; Length 300;
Best Local Similarity 35.0%; Pred. No. 33;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

QY 5 GTNGVHCTKNSLWCKVRLKNMKYDQNTTYMGRLODILLGW 44
DB 113 GTQVPCNKILLWSRIKDLAHQFTQVORDMFTLDTLLGY 152

RESULT 10
US-09-773-307B-2
; Sequence 2, Application US/09773307B
; Publication No. US20030027134A1
; GENERAL INFORMATION:
; APPLICANT: BML, INC.
; TITLE OF INVENTION: Method of Detecting Risk Factor for Onset of Diabetes
; FILE REFERENCE: PEM37
; CURRENT APPLICATION NUMBER: US/09/773,307B
; CURRENT FILING DATE: 2001-01-31
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Hominidae
US-09-773-307B-2

Query Match 19.7%; Score 60; DB 3; Length 300;





Qy 5 GTNGVHCTKNSLWKGKVLKNNKMYDQNTTYMGRLODILLGW 44  
Db 83 GTQVPCNKILLWSRIKDLAHQFTQVQORDMFTLEDTLLGY 122

Search completed: May 2, 2006, 17:45:42  
Job time : 165 secs

CURRENT APPLICATION NUMBER: US/10/753,267  
CURRENT FILING DATE: 2004-01-08  
PRIOR APPLICATION NUMBER: US 60/439,683  
PRIOR FILING DATE: 2003-01-13  
PRIOR APPLICATION NUMBER: US 60/445,216  
PRIOR FILING DATE: 2003-02-05  
PRIOR APPLICATION NUMBER: US 60/448,036  
PRIOR FILING DATE: 2003-02-18  
PRIOR APPLICATION NUMBER: US 60/454,189  
PRIOR FILING DATE: 2003-03-12  
PRIOR APPLICATION NUMBER: US 60/457,541  
PRIOR FILING DATE: 2003-03-25  
PRIOR APPLICATION NUMBER: US 60/466,411  
PRIOR FILING DATE: 2003-04-29  
PRIOR APPLICATION NUMBER: US 60/469,041  
PRIOR FILING DATE: 2003-05-08  
PRIOR APPLICATION NUMBER: US 60/477,414  
PRIOR FILING DATE: 2003-06-10  
PRIOR APPLICATION NUMBER: US 60/478,560  
PRIOR FILING DATE: 2003-06-13  
PRIOR APPLICATION NUMBER: US 60/489,772  
PRIOR FILING DATE: 2003-07-24  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 130  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 34  
LENGTH: 300  
TYPE: PRT  
ORGANISM: Homo Sapiens  
US-10-753-267-34

Query Match 19.7%; Score 60; DB 5; Length 300;  
Best Local Similarity 35.0%; Pred. No. 33;  
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy 5 GTNGVHCTKNSLWKGKVLKNNKMYDQNTTYMGRLODILLGW 44  
Db 113 GTQVPCNKILLWSRIKDLAHQFTQVQORDMFTLEDTLLGY 152

RESULT 15  
US-10-450-763-37242  
Sequence 37242, Application US/10450763  
Publication No. US20050196754A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc  
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
FILE REFERENCE: 790CIP3/US  
CURRENT APPLICATION NUMBER: US/10/450,763  
CURRENT FILING DATE: 2003-06-11  
PRIOR APPLICATION NUMBER: PCT/US01/08631  
PRIOR FILING DATE: 2001-03-30  
PRIOR APPLICATION NUMBER: 09/540,217  
PRIOR FILING DATE: 2000-03-31  
PRIOR APPLICATION NUMBER: 09/649,167  
PRIOR FILING DATE: 2000-08-23  
NUMBER OF SEQ ID NOS: 60736  
SOFTWARE: Custom  
SEQ ID NO 37242  
LENGTH: 688  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: DOMAIN  
LOCATION: (49)...(136)  
OTHER INFORMATION: ADP-ribosyl cyclase domain identified by Pfam, accession name  
OTHER INFORMATION: Rib\_hydrolyase, E-value=1.2e-88, Pfam score of 219.7  
US-10-450-763-37242

Query Match 19.7%; Score 60; DB 5; Length 688;  
Best Local Similarity 35.0%; Pred. No. 85;  
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

**THIS PAGE LEFT BLANK**

GenCore version 5.1.7  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: May 2, 2006, 17:43:06 ; Search time 25 Seconds  
(without alignments)  
98.232 Million cell updates/sec

Title: US-10-644-927-1  
Perfect score: 304  
Sequence: 1 KTYGTNGVHCTKNSLWGVK.....GRLQDILLGWATGAFGKTFF 54

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 232119 seqs, 45477862 residues  
Total number of hits satisfying chosen parameters: 232119

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA New:

- 1: /SIDSS5/prodata/2/pubpaa/US08\_NEW\_PUB.pep1.\*
- 2: /SIDSS5/prodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 3: /SIDSS5/prodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 4: /SIDSS5/prodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 5: /SIDSS5/prodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 6: /SIDSS5/prodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 7: /SIDSS5/prodata/2/pubpaa/US10\_NEW\_PUB.pep1.\*
- 8: /SIDSS5/prodata/2/pubpaa/US10\_NEW\_PUB.pep1.\*
- 9: /SIDSS5/prodata/2/pubpaa/US11\_NEW\_PUB.pep1.\*
- 10: /SIDSS5/prodata/2/pubpaa/US11\_NEW\_PUB.pep1.\*
- 11: /SIDSS5/prodata/2/pubpaa/US11\_NEW\_PUB.pep1.\*
- 12: /SIDSS5/prodata/2/pubpaa/US60\_NEW\_PUB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	66	21.7	557	11	US-11-188-298-21798
2	65.5	21.5	471	11	US-11-188-298-21598
3	62.5	20.6	446	11	US-11-188-298-13174
4	60	19.7	37	11	US-11-088-783-90
5	60	19.7	300	11	US-11-058-924-7
6	57.5	18.9	670	11	US-11-188-298-18003
7	57	18.8	473	11	US-11-188-298-15548
8	56.5	18.6	812	11	US-11-188-298-9586
9	56	18.4	127	11	US-11-188-298-13006
10	55.5	18.3	226	11	US-11-188-298-15017
11	54.5	17.9	480	9	US-10-915-002-299
12	53.5	17.6	56	9	US-10-921-415-3
13	53.5	17.6	134	9	US-10-921-415-4
14	53.5	17.6	278	11	US-11-087-099-587
15	53.5	17.6	353	11	US-11-087-099-5419
16	53.5	17.6	364	11	US-11-087-099-3112
17	53.5	17.6	373	11	US-11-087-099-7159
18	53.5	17.6	510	11	US-11-188-298-394
19	53.5	17.6	510	11	US-11-188-298-6950
20	53.5	17.6	510	11	US-11-188-298-21258
21	53.5	17.6	663	11	US-11-188-298-21626

Sequence 1, Appli  
Sequence 15, Appli  
Sequence 27618, A  
Sequence 27617, A  
Sequence 27616, A  
Sequence 13840, A  
Sequence 6622, Ap  
Sequence 247, App  
Sequence 7209, Ap  
Sequence 7731, Ap  
Sequence 4518, Ap  
Sequence 8664, Ap  
Sequence 16235, A  
Sequence 11478, A  
Sequence 6540, Ap  
Sequence 14116, A  
Sequence 10296, A  
Sequence 8332, Ap  
Sequence 916, App  
Sequence 22, Appl  
Sequence 6473, Ap  
Sequence 8800, Ap  
Sequence 15511, A  
Sequence 338, App

22 53.5 17.6 1033 9 US-10-921-415-1  
23 53.5 17.6 1033 9 US-10-501-841-15  
24 52.5 17.3 367 11 US-11-096-568A-27618  
25 52.3 17.3 385 11 US-11-096-568A-27617  
26 52.5 17.3 419 11 US-11-096-568A-27616  
27 52.5 17.3 510 11 US-11-188-298-13840  
28 52 17.1 193 11 US-11-079-463-6622  
29 52 17.1 719 9 US-10-511-538-247  
30 52 17.1 1263 11 US-11-087-099-7209  
31 51.5 16.9 381 11 US-11-079-463-7731  
32 51.5 16.9 472 11 US-11-188-298-4518  
33 51.5 16.9 510 11 US-11-188-298-8664  
34 51.5 16.9 510 11 US-11-188-298-16235  
35 51.5 16.9 511 11 US-11-188-298-11478  
36 51.5 16.9 513 11 US-11-188-298-6540  
37 51.5 16.9 513 11 US-11-188-298-14116  
38 51.5 16.9 1069 11 US-11-098-686-10296  
39 51 16.8 429 11 US-11-087-099-8332  
40 51 16.8 858 11 US-11-087-099-916  
41 51 16.8 3390 9 US-10-204-252-22  
42 50.5 16.6 355 11 US-11-079-463-6473  
43 50.5 16.6 399 11 US-11-079-463-8800  
44 50.5 16.6 444 11 US-11-188-298-15511  
45 50.5 16.6 491 9 US-10-506-454-338

## ALIGNMENTS

RESULT 1  
US-11-188-298-21798  
; Sequence 21798, Application US/11188298  
; Publication No. US20060075522A1  
; GENERAL INFORMATION:  
; APPLICANT: Abad, Mark S. et al.  
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-21(53452)B  
; CURRENT APPLICATION NUMBER: US/11/188,298  
; PRIOR FILING DATE: 2005-07-22  
; PRIOR APPLICATION NUMBER: 60/592,978  
; PRIOR FILING DATE: 2004-07-31  
; NUMBER OF SEQ ID NOS: 22569  
; SEQ ID NO 21798  
; LENGTH: 557  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(557)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-11-188-298-21798

Query Match 21.7%; Score 66; DB 11; Length 557;  
Best Local Similarity 36.5%; Pred. NO. 0.87; Mismatches 22; Indels 2; Gaps 2;  
Matches 19; Conservative 9;

QY 3 YYGINGVHCTKNSLW-GKVRLLKNMKYDQNT-TYMGRLQDILLGWATGAFGKT 52  
DB 66 YETTELVENKNGTYQWVVKPKTKVIEFKTNTHVPLGVMLVGVGXGNGNST 117

RESULT 2  
US-11-188-298-21598  
; Sequence 21598, Application US/11188298  
; Publication No. US20060075522A1  
; GENERAL INFORMATION:  
; APPLICANT: Abad, Mark S. et al.  
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT  
; FILE REFERENCE: 38-21(53452)B  
; CURRENT APPLICATION NUMBER: US/11/188,298  
; PRIOR FILING DATE: 2005-07-22  
; PRIOR APPLICATION NUMBER: 60/592,978  
; PRIOR FILING DATE: 2004-07-31

```
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 21598
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Sulfolobus solfataricus
US-11-188-298-21598

Query Match      21.5%; Score 65.5; DB 11; Length 471;
Best Local Similarity 48.6%; Pred. No. 0.84;
Matches 17; Conservative 2; Mismatches 13; Indels 3; Gaps 2;

Qy 1 KTYVGTN--GVHCTKNSLWGVRLKNMKYDQNTTY 33
Db 255 KPIYGDNGTGMH--THLSLWTKDGKKNLWYDPNDEY 288

RESULT 3
US-11-188-298-13174
; Sequence 13174, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 13174
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Methanococcus voltae
US-11-188-298-13174

Query Match      20.6%; Score 62.5; DB 11; Length 446;
Best Local Similarity 33.9%; Pred. No. 2;
Matches 19; Conservative 9; Mismatches 17; Indels 11; Gaps 5;

Qy 1 KTYVGTN--GVHCTKNSLW--GKVLKNMKYDQNTTYMGRLODILLGWATGAFGKT 52
Db 237 KPFGMGNGSHCNQ--SIWLDGRPSF---YDENNAH--QUSDICLSYIGILEHT 285

RESULT 4
US-11-068-783-90
; Sequence 90, Application US/11068783
; Publication No. US20050260715A1
; GENERAL INFORMATION:
; APPLICANT: Burian, Jan
; TITLE OF INVENTION: EFFICIENT METHODS FOR PRODUCING
; TITLE OF INVENTION: ANTIMICROBIAL CATIONIC PEPTIDES IN HOST CELLS
; FILE REFERENCE: 660081.411
; CURRENT APPLICATION NUMBER: US/11/068,783
; CURRENT FILING DATE: 2005-02-28
; PRIOR APPLICATION NUMBER: US/09/444,281
; PRIOR FILING DATE: 1999-11-19
; NUMBER OF SEQ ID NOS: 113
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 90
; LENGTH: 37
; TYPE: PRT
; ORGANISM: Leuconostoc gelidum
US-11-068-783-90

Query Match      19.7%; Score 60; DB 11; Length 37;
Best Local Similarity 48.4%; Pred. No. 0.26;
Matches 15; Conservative 2; Mismatches 4; Indels 10; Gaps 3;

Qy 3 YVYNGVHCTKNSL---WGKV-----RLKN 24
Db 2 YVY-NGVHCTKSGCSVNWGEAFSGVHRLAN 31
```

```
RESULT 5
US-11-058-924-7
; Sequence 7, Application US/11058924
; Publication No. US20060019308A1
; GENERAL INFORMATION:
; APPLICANT: Lund, Frances E.
; APPLICANT: Randall, Troy D.
; TITLE OF INVENTION: CD38 MODULATED CHEMOTAXIS
; FILE REFERENCE: 13315/5
; CURRENT APPLICATION NUMBER: US/11/058,924
; CURRENT FILING DATE: 2005-02-15
; PRIOR APPLICATION NUMBER: 09/982,616
; PRIOR FILING DATE: 2001-10-17
; PRIOR APPLICATION NUMBER: 60/241,065
; PRIOR FILING DATE: 2000-10-17
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 7
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-058-924-7

Query Match      19.7%; Score 60; DB 11; Length 300;
Best Local Similarity 35.0%; Pred. No. 2.7;
Matches 14; Conservative 7; Mismatches 19; Indels 0; Gaps 0;

Qy 5 GTNGVHCTKNSLWGVRLKNMKYDQNTTYMGRLODILLGW 44
Db 113 GTQVPCNKILLWSRIKDLAHOFTQVQDMFTLEDLLGY 152

RESULT 6
US-11-188-298-18003
; Sequence 18003, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 18003
; LENGTH: 670
; TYPE: PRT
; ORGANISM: GIBBERELLA ZEAE PH-1
US-11-188-298-18003

Query Match      18.9%; Score 57.5; DB 11; Length 670;
Best Local Similarity 29.5%; Pred. No. 14;
Matches 13; Conservative 8; Mismatches 18; Indels 5; Gaps 2;

Qy 3 YVG--TNGVHCTKNSLWGVRLKNMKYDQNTTYMGRLODILLGW 44
Db 566 YHGVAVNPJHL---SLWERLRLGLGKSYDPEADWRKAKVQDMRAEW 606

RESULT 7
US-11-188-298-15548
; Sequence 15548, Application US/11188298
; Publication No. US20060075522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
```

```
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 15548
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Sulfolobus acidocaldarius
US-11-188-298-15548

Query Match      18.8%; Score 57; DB 11; Length 473;
Best Local Similarity 45.7%; Pred. No. 11;
Matches 16; Conservative 4; Mismatches 11; Indels 4; Gaps 3;

Qy 1 KTYGTN--GVHCTKNSLWGVRLKNNKYDNTTY 33
|::|||::|||::|||::|||::|||::|||::|||
Db 257 KPFFDNGSGMH--THFSLWTXDG-KNLWYDPNDEY 289

RESULT 8
US-11-188-298-9586
; Sequence 9586, Application US/11188298
; Publication No. US2006007522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 9586
; LENGTH: 812
; TYPE: PRT
; ORGANISM: Eremothecium gossypii
US-11-188-298-9586

Query Match      18.6%; Score 56.5; DB 11; Length 812;
Best Local Similarity 29.4%; Pred. No. 24;
Matches 15; Conservative 7; Mismatches 18; Indels 11; Gaps 1;

Qy 4 YGTNGVHCTKNSLWGVRLKNNKYDNTTYMGRLODILLGWATGAFGKTFH 54
|::|||::|||::|||::|||::|||::|||::|||
Db 331 YGT-----GRAVLFEVSYQSNWRYLEIVLVILGAAGVYGALFN 370

RESULT 9
US-11-188-298-13006
; Sequence 13006, Application US/11188298
; Publication No. US2006007522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 13006
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1), (147)
; OTHER INFORMATION: unsure at all xaa locations
US-11-188-298-13006

Query Match      18.4%; Score 56; DB 11; Length 147;
Best Local Similarity 32.7%; Pred. No. 4.2;
Matches 17; Conservative 10; Mismatches 23; Indels 2; Gaps 2;

Qy 3 YYGTNGVHCTKNSLW-GKVRLLKNNKYDQNT-TYMGRLQDILLGWATGAFGKT 52
```

```
Db 26 YETTELVHENRNGTYQWIVKPKSVNYQFKTNTHTVPLKGLVMLVGVGXGNGST 77

RESULT 10
US-11-188-298-15017
; Sequence 15017, Application US/11188298
; Publication No. US2006007522A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: GENES AND USES FOR PLANT IMPROVEMENT
; FILE REFERENCE: 38-21(53452)B
; CURRENT APPLICATION NUMBER: US/11/188,298
; CURRENT FILING DATE: 2005-07-22
; PRIOR APPLICATION NUMBER: 60/592,978
; PRIOR FILING DATE: 2004-07-31
; NUMBER OF SEQ ID NOS: 22569
; SEQ ID NO 15017
; LENGTH: 226
; TYPE: PRT
; ORGANISM: Glycine max
US-11-188-298-15017

Query Match      18.3%; Score 55.5; DB 11; Length 226;
Best Local Similarity 36.5%; Pred. No. 7.9;
Matches 19; Conservative 9; Mismatches 21; Indels 3; Gaps 3;

Qy 3 YYGTNGVHCTKNSLW-GKVRLLKNNKYDQNT-TYMGRLQDILLGWATGAFGKT 52
|::|||::|||::|||::|||::|||::|||::|||
Db 26 YETTELVHENRNGTYQWIVKPKSVNYQFKTNTHTVPLKGLVMLVGVGXGNGST 76

RESULT 11
US-10-915-002-299
; Sequence 299, Application US/10915002
; Publication No. US20060078950A1
; GENERAL INFORMATION:
; APPLICANT: Proguleske-Fox, Ann
; APPLICANT: Hillman, Jeffrey D.
; APPLICANT: Handfield, Martin
; TITLE OF INVENTION: IDENTIFICATION OF PORPHYROMONAS GINGIVALIS VIRULENCE POLYNUCLEOTI
; TITLE OF INVENTION: USE IN DIAGNOSIS ANTIGENS FOR USE IN THE DIAGNOSIS, TREATMENT, A
; TITLE OF INVENTION: PERIODONTAL DISEASES
; FILE REFERENCE: 02-042
; CURRENT APPLICATION NUMBER: US/10/915,002
; CURRENT FILING DATE: 2004-08-10
; NUMBER OF SEQ ID NOS: 354
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 299
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Porphyromonas gingivalis
US-10-915-002-299

Query Match      17.9%; Score 54.5; DB 9; Length 480;
Best Local Similarity 28.2%; Pred. No. 25;
Matches 11; Conservative 12; Mismatches 13; Indels 3; Gaps 2;

Qy 3 YYGTNGVH--CTKNSL-WGKVRLLKNNKYDQNTTYMGRLO 38
|::|||::|||::|||::|||::|||::|||::|||
Db 296 FYAKSGMHFYATKHNFEWDKKSIGKMLFDRDKDLEGKMK 334

RESULT 12
US-10-921-415-3
; Sequence 3, Application US/10921415
; Publication No. US20060014681A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Xiaojiang
; APPLICANT: Holers, V. Michael
; TITLE OF INVENTION: THREE-DIMENSIONAL STRUCTURE OF COMPLEMENT RECEPTOR TYPE 2 AND US
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 2848-43
```

```

; CURRENT APPLICATION NUMBER: US/10/921,415
; CURRENT FILING DATE: 2004-08-16
; PRIOR APPLICATION NUMBER: US/09/834,309
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-921-415-3

```

Query Match 17.6%; Score 53.5; DB 9; Length 56;  
Best Local Similarity 44.0%; Pred. NO.3.1;  
Matches 11; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

Qy 1 KTYYGTTNG---VHCTKNSLWGKVRL 22  
|||: ||| |||: ||| |||  
Db 29 KTNFSMNGNKSVWCQANNMWGPTRL 53

```

RESULT 13
US-10-921-415-4
; Sequence 4, Application US/10921415
; Publication No. US20060014681A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Xiaojiang
; APPLICANT: Holers, V. Michael
; TITLE OF INVENTION: THREE-DIMENSIONAL
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 2848-43
; CURRENT APPLICATION NUMBER: US/10/
; CURRENT FILING DATE: 2004-08-16
; PRIOR APPLICATION NUMBER: US/09/83
; PRIOR FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 134
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-921-415-4

```

Query Match 17.6%; Score 53.5; DB 9; Length 134;  
Best Local Similarity 44.0%; Pred. No. 8.1;  
Matches 11; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

Qy 1 KTYYG TNG-- --VHCTKNSLW GKVRL 22  
||| : ||| | | : ||| ||  
Db 100 KTNFMSMGNKSVWCQANNMWGPTRL 124

```

RESULT 14
US-11-087-099-587
; Sequence 587, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B Ep
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 587
; LENGTH: 278
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(278)
; OTHER INFORMATION: unsure at all Xaa locations
US-11-087-099-587

```

	Query March	17.6%	Score 53.5;	DB 11;	Length 278;
Best Local Similarity	28.6%				
Matches 10;	Conservative 8;	Pred. No. 18;			
				Indels 1	
Qy	5 GTNGVHCITKSLGWKVRLLKNMKYDQNTTYMGLQD	39			
Db	128 GHEGWHC-OGRRWGGIMKEKRNPKRGSSYMGSIYD	161			

```

RESULT 15
US-11-087-099-5419
; Sequence 5419, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450)B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 5419
; LENGTH: 353
; TYPE: PRP
; ORGANISM: Mesembryanthemum crystallinum
US-11-087-099-5419

```

```

Query Match      17.6%; Score 53.5; DB 11; Length 353;
Best Local Similarity 41.4%;
Matches 12; Conservative 4; Mismatches 8; Indels 5

Qy      24 NMKYDQNTTYNGRLQDIL-----LGWATG 47
      : : : : : : : : : : : : : : : :
Db      177 NIETMRNTLYKAYLEDYRFFCOKLGWATG 205

```

Search completed: May 2, 2006, 17:46:14  
Job time : 26 secs